Verwendung von Geolokationsdaten als Angriffsvektor für Social Engineering

Matteo Cagnazzo · Norbert Pohlmann

Institut für Internet-Sicherheit – if(is)
Westfälische Hochschule
{cagnazzo | pohlmann}@internet-sicherheit.de

Zusammenfassung

Geolokationsdaten sind eine Bedrohung für Organisationen und Unternehmen. Vor kurzem wurde beispielsweise ein anonymisierter Datensatz veröffentlicht, allerdings sind die verwendeten Anonymisierungstechniken unzureichend. Private und eventuell geheime Standorte, beispielsweise Militärbasen in Kriegsgebieten, können damit aufgedeckt werden. Weiterhin können einzelne Personen gezielt identifiziert werden.

1 Einführung

Frei zugängliche Daten von privaten Firmen sind eine wichtige Quelle für unterschiedliche Forschungsbereiche. Es gibt viele Beispiele für die positive Nutzung von öffentlichen Daten in der Forschung zum Beispiel im Bereich des nicht-motorisierten Transportes oder Untersuchungen zur Exposition in Gebieten mit extremer Luftverschmutzung [SELA16] [SUN17].

Die aktuellen Entwicklungen und Erkenntnisse zeigen aber auch deutlich Probleme und Gefahren auf, die mit der Veröffentlichung solcher Datensätze einhergehen. Insbesondere die Betriebssicherheit (Opsec) und Privatheit von Individuen und großen Organisationen kann durch die Veröffentlichung solcher Daten gefährdet sein. Dieses Paper zielt darauf ab Gefahren aufzuzeigen und Awareness zu schaffen, dass solche Probleme existieren. Weiterhin werden mögliche Konsequenzen von betrieblichen- oder Privatsphäre Problemen angerissen und diskutiert. Das Kapitel 2 gibt Hintergrundinformationen für den Leser in den Bereichen Social Engineering, Privatsphäre und ortsbasierte Daten. Kapitel 3 zeigt Ergebnisse einer kleinen Fallstudie auf, welche durchgeführt wurde, um die Machbarkeit zu zeigen. Anschließend werden in Kapitel 4 mögliche Mitigierungsstrategien diskutiert. Im letzten Kapitel 5 werden die Ergebnisse diskutiert und weitere Forschungsfragen skizziert.

2 Hintergrundinformationen

Dieses Kapitel gibt einige theoretische Definitionen und Einführungen für ortsbasierte Daten, Social Engineering und Angriffstechnologien.

P. Schartner · N. Pohlmann (Hrsg.) · D•A•CH Security 2018 · syssec (2018) 383-391.

2.1 Ortsbasierte Daten

Ortsbasierten Daten enthalten Ortsinformationen in Form von Koordinaten und üblicherweise einen Zeitstempel. Diese Daten werden genutzt, um beispielsweise Geräte, Personen oder andere Entitäten zu orten. Das meistgenutzte System ist das Global Positioning System. Dieses arbeitet mit der Triangulation von Radiosignalen und Satelliten. [MISR06] gibt eine ausführliche Erläuterung und Überblick über diese Technologie. GPS wird in vielen Smartphones, mobilen Geräten oder Dingen im Internet der Dinge implementiert. Durch die ubiquitäre Vernetzung dieser Gegenstände wird es möglich, das Gerät jederzeit zu orten oder Informationen über den derzeitigen Aufenthaltsort zu versenden.

Diese Informationen werden von Applikationen genutzt, um zum Beispiel Laufstreckentracking zu ermöglichen. Ein großer Dienst in diesem Bereich ist das soziale Netzwerk Strava. Strava ist eine populäre Applikation, welche von Athleten jeglicher Leistungsklasse genutzt wird, um sportliche Aktivitäten mit anderen Sportlern zu vergleichen. Mit Hilfe von Applikationen wie Strava können Parameter wie zum Beispiel Laufdistanz, Laufzeit, Durchschnittsgeschwindigkeit, und Route mit anderen Nutzern geteilt werden. Dabei gibt es auch die Möglichkeit viele Metainformationen anzugeben, beispielsweise welches Laufschuhmodell oder welche Fitnessarmbanduhr wurde verwendet. Dienste wie Strava oder Suunto veröffentlichen anonymisierte Heatmaps der Aktivitäten ihrer Nutzer öffentlich zugänglich. Diese Daten wurden dann misbraucht, um militärische Stellungen überall auf der Welt, auch in Krisenregionen, aufzudecken [GUAR18]. Die Heatmap besteht aus drei Trillionen ortsbezogenen Daten in aggregierter Form. Die zeitlichen Komponenten der Daten wurden ebenfalls entfernt. Insgesamt hat allein Strava nach eigenen Angaben eine Milliarde Aktivitäten und zehn Millionen Nutzer in ihren Daten [STRA17].

Innerhalb der Fallstudie wird klar, wie schnell einzelne Nutzer anhand der Daten identifiziert werden können, beispielsweise als Soldaten auf einem Stützpunkt oder Arbeiter bestimmter Firmen. Durch die Kombination der Heatmap und den sogenannten „Segmenten“ ist dies möglich. Segmente sind „user-created-content“, das bedeutet Nutzer können Segmente anlegen um mit anderen Mitgliedern in einen virtuellen Wettstreit zu treten. Dafür wird nach dem absolvieren eines Laufes ein Streckenabschnitt innerhalb der Applikation markiert und benannt. Die Applikation erstellt nun eine Bestenliste für die einzelnen Segmente. Durch den virtuellen Wettstreit sollen die Nutzer der Applikation motiviert bleiben während sie joggen oder Rad fahren. Diese Segmente sind in der Standardeinstellung öffentlich. Man kann die Segmente nur für einen kleinen Personenkreis sichtbar machen, allerdings ist dann eines der „Key-Features“ der Applikation nicht funktional, der „virtuelle Wettstreit“ und die soziale Teilhabe an dem Erfolg des anderen.

2.2 Social Engineering

Social Engineering ist im Informationssicherheitskontext die Beeinflussung von potentiellen Opfern Informationen preiszugeben oder Handlungen durchzuführen durch soziale Manipulation. Diese Informationen oder Handlungen werden dann dafür genutzt ein informationsverarbeitendes System oder eines anderen Assets innerhalb eines Unternehmens zu manipulieren. Die Manipulation kann sich auf die Vertrauenswürdigkeit, Integrität oder die Verfügbarkeit auswirken. Täuschung und Manipulation sind die Grundlagen von Social Engineering-gestützten Angriffen.

Aktuelle technische Gegenmaßnahmen sind in der Regel gegen diese Angriffe unwirksam. Darüber hinaus haben viele Opfer von Social Engineering Angriffen die Meinung, dass sie gut darin sind, diese Angriffe zu erkennen, es aber eigentlich nicht sind [KROM15]. Es gibt mehrere Stufen innerhalb der Taxonomie eines erfolgreichen Social Engineering-Angriffs, wie in Abbildung 1 zu sehen ist. Vor allem wenn der Angriff darauf abzielt, ein Unintentional Insider Threat (UIT) zu beinhalten. [BURE13] definiert einen UIT wie folgt: „Ein Unintentional Insider Threat ist (1) ein gegenwärtiger oder ehemaliger Mitarbeiter, Auftragnehmer oder Geschäftspartner (2), der Zugriff auf das Netzwerk, System oder Daten einer Organisation autorisiert hat oder hatte und (3) durch Aktion oder Unterlassung ohne böswillige Absicht (4) verursacht Schaden oder erhöht wesentlich die Wahrscheinlichkeit eines zukünftigen ernsthaften Schadens für die Vertraulichkeit, Integrität oder Verfügbarkeit des Informations- oder Informationssystems der Organisation“.

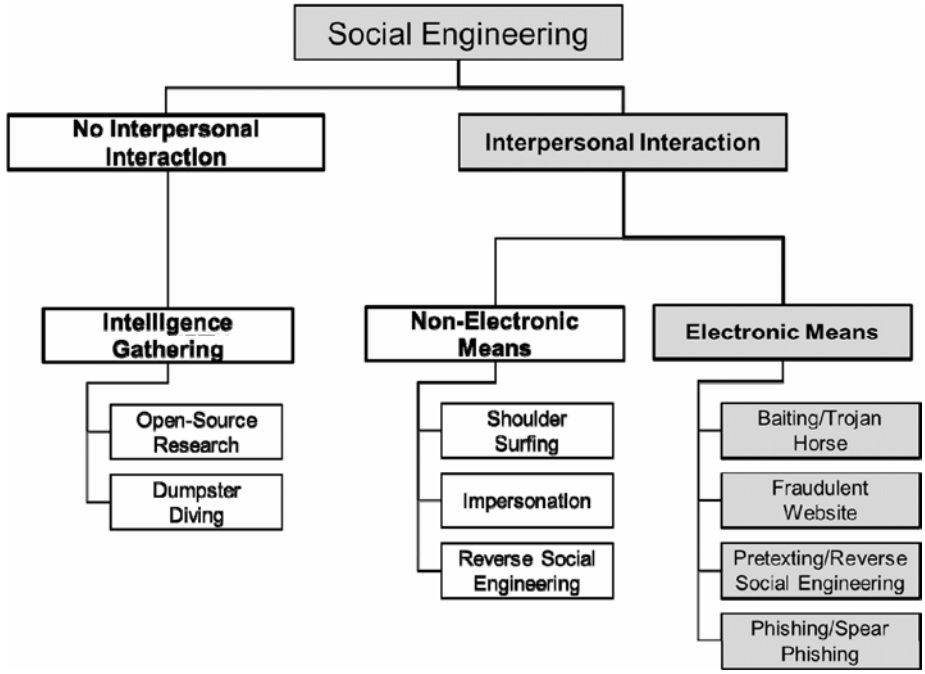

Abb. 1: Taxonomie eines Social Engineering Angriffs [GREI14]

Diese Insider-Bedrohungen, unabhängig davon ob beabsichtigt oder nicht, gelten als eines der größten Risiken für die betriebliche und organisatorische Sicherheit. Die beiden Bedrohungen unterscheiden sich in Bezug auf Motivation, Indikation und andere Unterschiede. Daher ist es wichtig, diese Bedrohungen zu erforschen und ihr Aufkommen zu verstehen [BURE13]. Für die von uns durchgeführte Fallstudie werden wir Open Source Informationen von Personen mit standortbezogenen Informationen aus Fitnessapplikationen verknüpfen. Aus den gesammelten Informationen lassen sich gezielte Pretexte und zielgruppenspezifisches (Spear) Phishing entwickeln.

Beim Pretexting wird ein erfundenes oder echtes Szenario verwendet, um die Chancen zu erhöhen, dass ein Opfer die Vertraulichkeit, Integrität oder Verfügbarkeit seines Unternehmens gefährdet. Unter normalen Bedingungen würde das Opfer eine solche Handlung nicht durchführen, aber durch die geschickte Wahl eines Pretext kann ein potentielles Opfer motiviert werden, auf bösartige Links oder Anhänge zu klicken.

Spear Phishing enthält häufig Pretexting die den „Phisher“ als vertrauenswürdige dritte Partei darstellen. Das erhöht die Erfolgswahrscheinlichkeit eines Angriffs, selbst wenn Benutzer geschult sind und eine gesteigerte Awareness besitzen [CAPU14]. Advanced Persistent Threats können Spear Phishing als initialen Kompromittierungsvektor enthalten, wie zum Beispiel „Operation Pawn Storm“ oder „TG-4127“ [SECW18] [TREN15].

2.3 Angreifer Modelle

Wir unterscheiden zwischen zwei Angreifer Modellen. Das erste Modell ist ein passiver Angreifer, der nur Informationen lesen kann und keine Daten auf der Plattform verändert. Wir können weiter zwischen einem angemeldeten Nutzer unterscheiden, der mehr Informationen sammeln kann und einem Besucher, der eingeschränkt Daten sehen kann. Der zweite Angreifer ist ein aktiver Angreifer. Er lädt speziell gestaltete Daten oder Texte hoch, um Benutzer unter bestimmten Umständen zu identifizieren.

3 Fallstudie

In diesem Abschnitt werden die Ergebnisse der Fallstudie veröffentlicht und gezeigt, wie aus Daten in Fitnessnetzwerken Merkmale zur Identifizierung extrahiert werden können. Diese detaillierten Informationen lassen sich dann für einen individualisierten Pretext oder eine individualisierte Phishing Kampagne nutzen.

3.1 Heatmaps und Segmente

Nachdem Strava die Heatmap veröffentlicht hat, haben wir eine Fallstudie durchgeführt. Das Angreifermodell, welches wir verwenden, ist ein passiver Angreifer, der ein eingeloggter User ist. Um vollwertiger User auf der Plattform zu werden reicht eine beliebige E-Mail-Adresse und ein Passwort zur Registrierung. Im ersten Schritt werden zufällig zwanzig Militärbasen auf der ganzen Welt ausgewählt und nach Segmenten auf oder in direkter Umgebung der Basen gesucht. Liegen die Segmente innerhalb eines militärischen Sperrgebiets kann davon ausgegangen werden, dass die gefundenen User im Dienst stehen.

Oft sind Nutzer bei Strava mit ihrem Klarnamen angemeldet. Die Nutzung eines Pseudonyms ist eher die Ausnahme. Liegen die Segmente in der Umgebung der Basis, wird die Identität der Nutzer durch Zeitungsartikel oder andere soziale Netzwerke, beispielsweise LinkedIn verifiziert. Von den zwanzig untersuchten Basen liegen in zwei Basen keine Segmente auf militärischem Gebiet. In jeder der Basen war es uns möglich mindestens einen Nutzer zu identifizieren und die Identität anhand von den oben aufgeführten Querverweisen zu verifizieren.

Beispielsweise hat eine Panzerkaserne in Böblingen eine verdächtige Heatmap-Struktur, in Form einer klassischen Laufbahn, wie in Abbildung 2 zu sehen ist. Um dies zu verifizieren, wird eine klassische Karte, beispielsweise Open Street Map oder Google Maps geöffnet. Es ist nun zu sehen, dass dort die „Stuttgart High School“ auf der Panzerkaserne ist. Dies ist in Abbildung 3 zu sehen. Abbildung 3 zeigt außerdem das Segment in der Strava App, welches mit „D“ markiert ist. Dieses Segment ist häufig frequentiert und zum Zeitpunkt der Studie ließen sich mehr als 20 amerikanische Soldaten, die in der Panzerkaserne in Böblingen stationiert waren, identifizieren. Dieses Segment wurde mittlerweile entfernt und ist nicht mehr einsehbar.

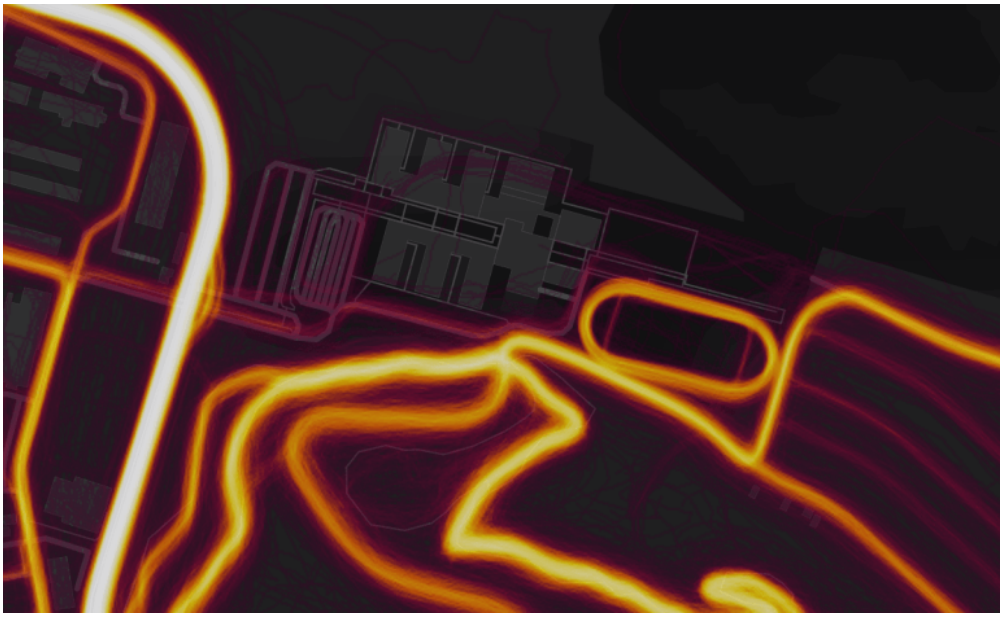

Abb. 2: Heatmap Signatur der Laufbahn in Böblingen

Jeder Benutzer, der dieses Segment lief und seine Leistungen mit anderen verglichen hat, war so sichtbar. Nicht überraschend ist die Tatsache, dass auf militärischem Gelände nahezu ausschließlich Soldaten laufen. Auch die entsprechenden Profile in den Applikationen zeigen eine enge Verbindung zum Militär, beispielsweise durch die Gruppenzugehörigkeiten in „Royal British Airforce Running Veterans“.

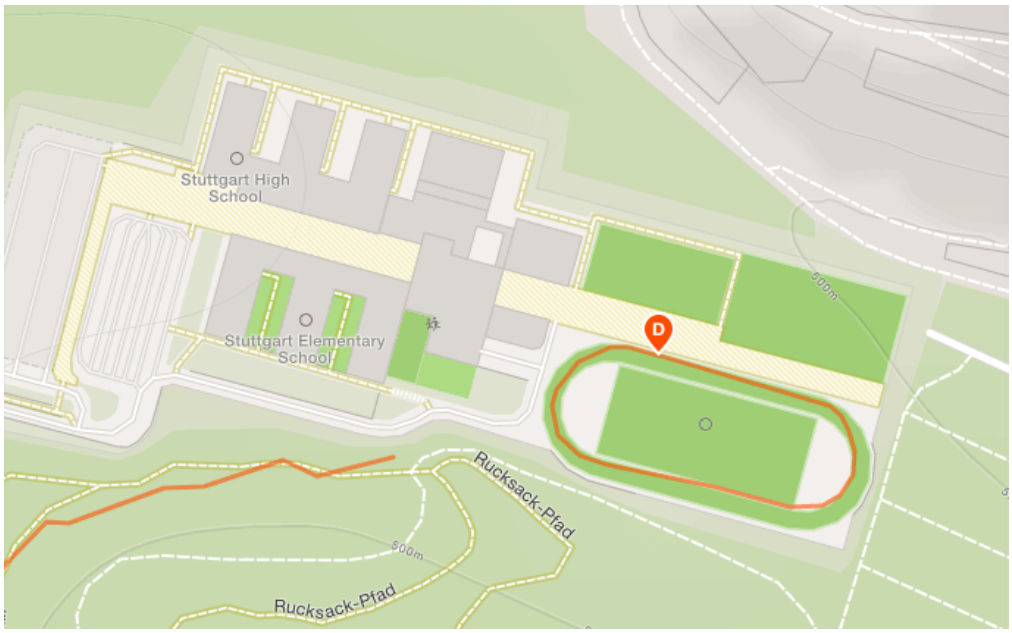

Abb. 3: Karte der fraglichen Region

Ein weiteres interessantes Ergebnis ist die Tatsache, dass identifizierte Nutzer über die Zeit hinweg verfolgt werden können. Auf dem Profil der Nutzer finden sich beispielsweise auch andere Basen, auf denen der Nutzer stationiert war.

Auf Abbildung 4 ist beispielsweise die Laufroute eines eigentlich in Böblingen stationierten Soldaten zu sehen, der irgendwann in das US Naval Diving & Salvage Training Center in Panama City Beach, Florida umstationiert wurde.

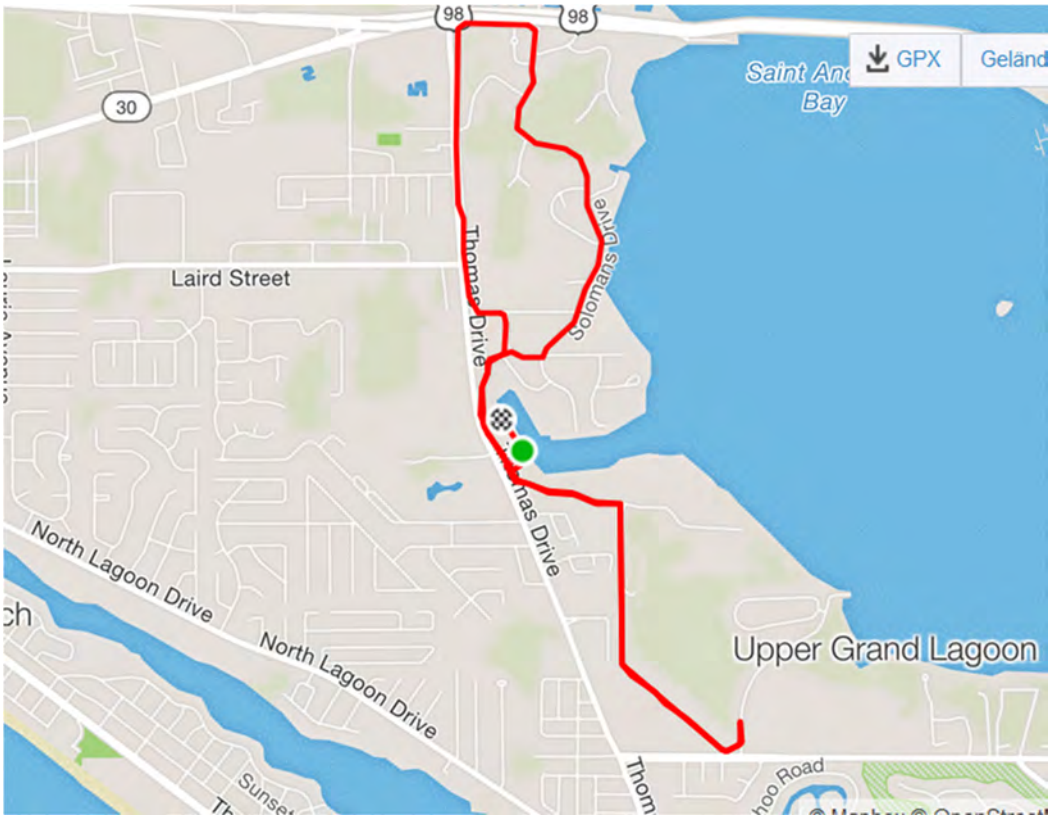

Abb. 4: Umstationierung eines Nutzers

Innerhalb der zwanzig Basen konnten sechs Personen zu mindestens einer weiteren Militärbasis oder ihren Heimatort verfolgt werden. Es ist möglich die Heimatstädte durch andere soziale Netzwerke, beispielsweise Facebook, zu verifizieren. Auch das Feststellen von Urlaubsphasen ist möglich. Es gibt innerhalb von Strava Datenschutzeinstellungen, die ein solches Informationsleck schließen können aber sie sind vielschichtig und standardmäßig sind diese Einstellungen ein „opt-out“. Während dieser Fallstudie haben wir keine automatisierten Verfahren entwickelt. Ein vollautomatisiertes scraping von Segmenten war zu diesem Zeitpunkt allerdings möglich: Segmente haben sich zum Zeitpunkt der Fallstudie hinter folgendem Schema befunden: „www.strava.com/segments/########“. Das „#“ steht dabei für eine Ziffer. Diese Struktur wurde seitens Strava aufgegeben aufgrund der leichten Enumeration und der Anfälligkeit für die vollautomatisierte Auswertung der Segmente durch Dritte. Um eine Vorschau auf die Rangliste des Segments mit bis zu 40 Athleten (Die 20 besten insgesamt und 20 Frauen/Männer) zu sehen kann sogar ein passiver Angreifer, der nicht eingeloggt ist, die Daten auslesen. Die Terms of Service verbieten dies, aber da Kriminelle bereits das Gesetz brechen, werden Terms of Service Sie nicht aufhalten. Wenn sich ein Segment also innerhalb eines Militärstandorts oder in naher Umgebung befindet, ist die Wahrscheinlichkeit hoch, dass sich stationierte Personen identifizieren lassen.

Abhängig von den individuellen Datenschutzeinstellungen der Nutzer lässt sich zudem bestimmen, welche Art Fitness-Tracker, Laufschuh oder Fahrrad ein Benutzer verwendet.

Aus Angreifersicht sind dies wichtige Informationen, da er nun Details kennt über die exakte athletische Ausrüstung. Der Angreifer könnte nun beispielsweise Pretexte oder Phishing Mails auf exakt dieses Produkt zuschneiden.

Wechseln wir nun in den Unternehmensbereich, sind dort alle oben genannten Bedrohungen ebenfalls vorhanden. Insbesondere ein aktiver Angreifer kann viele Informationen über Mitarbeiter von Unternehmen sammeln. Es lassen sich bestimmte Punkte auf der Heatmap identifizieren von der scheinbar viele Nutzer starten, wie in Abbildung 5 zu sehen ist.

Abb. 5: Startpunkte an einem Unternehmen

Diese Startpunkte befinden sich auf einem Firmengelände. Wenn ein Angreifer diese Punkte kennt, muss er eine gültige .gpx-Datei erzeugen, die diese Startpunkte durchläuft. Wenn ein Angreifer physikalisch in der Nähe ist, kann er einfach diese Route laufen oder Radfahren und die Aktivität hochladen. Möchte ein Angreifer beliebige Informationen, unabhängig davon wo er oder das Unternehmen geographisch ist, sammeln kann er eine .gpx Datei hochladen und eine Aktivität hinzufügen. Das Schema der .gpx-Datei ist wie folgt:

<trkpt lat="51.#####" lon="-6.#####"/> <time>2018-02-28T09:32:21Z</time>

Manche Dienste ermöglichen es nun auf Strecken gegen andere Benutzer zu konkurrieren virtuell. Läuft ein Benutzer ein 1,5 km langes Segment in 5:00 muss ein anderer Benutzer die Zeit unterbieten um auf dem Leaderboard vor den Ersteller zu kommen. Diese Leaderboards sind unter den oben angegebenen Einschränkungen zugänglich.

Durch das hinzufügen einer solchen Route, die ein kurzes Segment im Umkreis der Startpunkte in Abbildung 5 hat, konnten innerhalb kurzer Zeit mehrere Mitarbeiter aus der beobachteten Organisation identifiziert werden. Sie konkurrierten unfreiwillig mit unserer Zeit und wurden dadurch dem Leaderboard hinzugefügt. Dies führt zu einer granularen Identifizierung mit nützlichen Informationen für Social Engineering Kampagnen. Die Tatsache, dass Dienste wie Strava noch mehr Informationen liefern als den Namen ist dabei ein Vorteil auf Angreiferseite. Über ein solches Segment lässt sich ableiten, ob ein Nutzer beispielsweise regelmäßig mit dem

Fahrrad zur Arbeit pendelt oder nach der Arbeit oder in der Mittagspause laufen geht. Zusätzlich stehen ebenfalls Metainformationen wie beispielsweise Informationen über den Tracker und Laufschuh zur Verfügung. All dies sind wichtige Informationen für Angreifer, die in gezielten Kampagnen gegen einzelne Unternehmen vorgehen und versuchen den perfekten initialen Angriffsvektor zu finden.

4 Mögliche Mitigierungsstrategien

Diese Risiken zu minimieren ist schwierig. Soziale Netzwerke haben häufig mit inkohärenten Privatsphäre Einstellungen zu kämpfen. Insbesondere bei Fitnessapplikationen wird fehlende Transparenz kontrovers diskutiert [SPIN17]. Die Heatmaps die durch Dienste veröffentlicht werden, könnten „privacy-by-design“ Ansätze implementieren, so wie in [OKSA15] beschrieben. Eine weitere Minimierung könnte ein höheres Bewusstsein der Benutzer sein, bei der Verwendung von Geolokalisations-basierten Applikationen. Es gibt Möglichkeiten sich von dem Erscheinen in Segmenten und Rankings auszuschließen, aber die meisten Nutzer verwenden diese Einstellungen nicht, da Ihnen die Partizipation wichtiger erscheint, wenn sie eine Aktivität hochladen. Auch das Bewusstsein für aktuelle Social Engineering Techniken könnte ein Faktor sein, der zur Resilienz gegen solche Angriffe beiträgt.

5 Resümee

Zusammenfassend veranschaulicht unsere Fallstudie wie vernetzt Nutzer solcher Applikationen mittlerweile sind und wie granular Daten über diese Nutzer gesammelt werden. Die Bedrohungen, die dadurch entstehen sind mannigfaltig, insbesondere da durch solche Technologien die Grenze zwischen Privat- und Arbeitsleben immer weiter verschwindet. Durch dieses Verwischen der Abgrenzung können Nutzer leicht zum UIT innerhalb ihres Unternehmen werden. Standortdaten bieten für viele Parteien also potentielle Risiken, nicht nur für die Privatperson, welche die Applikation als Fitness-Tracker nutzt. Die hohe Granularität der Heatmap und die Leichtigkeit mit der Personen in Sicherheitszonen identifiziert werden können unterstreicht die Probleme bei geolokationsbasierten Applikationen. Besonders in Ländern mit hohem militärischem Engagement, wie beispielsweise Israel oder den USA ist dies eine Gefahr für die Sicherheit und Privatsphäre von Einzelpersonen und ganzen Organisationen. Eine privat genutzte Sportapplikation kann eine Bedrohung für die organisatorische Sicherheit, Integrität, Vertraulichkeit und Verfügbarkeit sein.

Literatur

[BURE13] Bureau, Federal Infrastructure Protection. Unintentional Insider Threats: A Foundational Study (2013).

[CAPU14] D. D. Caputo et al.: Going spear phishing: Exploring embedded training and awareness. IEEE Security & Privacy 12.1 (2014) 28-38.

[GUAR18] The Guardian, (2018) Fitness tracking app Strava gives away location of secret US army bases. https://www.theguardian.com/world/2018 /jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases Accessed Feb. 14 2018.

[GREI14] F. L. Greitzer et al.: Analysis of unintentional insider threats deriving from social engineering exploits. IEEE Security and Privacy Workshops (2014).

[KROM15] K. Krombholz et al.: Advanced social engineering attacks. Journal of Information Security and applications 22 (2015) 113-122.

[MISR06] P. Misra, P. Enge: Global positioning system: Signals, measurements and performance, 2nd edition. Ganga-Jamuna Press (2006).

[OKSA15] J. Oksanen et al: Methods for deriving and calibrating privacy-preserving heat maps from mobile sports tracking application data. Journal of Transport Geography 48 (2015) 135-144.

[SECW18] Secure Works: Threat Group-4127 Targets Hillary Clinton Presidential Campaign. https://www.secureworks.com/research/threat-group-4127-targets-hillary-clinton-presidential-campaign. Accessed Feb. 14 2018]

[SELA16] M. K. Selala, W. Musakwa: The potential of strava data to contribute in non-motorised transport (Nmt) planning in Johannesburg. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 41 (2016) 587.

[SPIN17] R. Spinks: Using a fitness app taught me the scary truth about why privacy settings are a feminist issue. Aug. 01, 2017. https://qz.com/1042852/using-a-fitness-app-taught-me-the-scary-truth-about-why-privacy-settings-are-a-feminist-issue/. Accessed Mar. 30 2018.

[STRA17] Strava stories, 2017 in Stats. https://blog.strava.com/2017-in-stats/ Accessed Feb. 14 2018.

[SUN17] Y. Sun, A. Mobasheri: Utilizing Crowdsourced data for studies of cycling and air pollution exposure: A case study using Strava Data. International journal of environmental research and public health 14.3 (2017) 274.

[TREN15] Trend Micro (2015) Operation Pawn Storm Ramps Up its Activities; Targets NATO, White House. https://blog.trendmicro.com/trendlabs- security-intelligence /operation- pawn- storm-ramps-up-its-activities-targets-nato-white-house/. Accessed Feb. 14 2018.