Data Breaches in IT Systems

Magdalena Neumann
Betreuer: Heiko Niedermayer
Seminar: Innovative Internettechnologien und Mobilkommunikation SS2016
Lehrstuhl Netzarchitekturen und Netzdienste
Fakultät für Informatik, Technische Universität München
Email: neumamag@in.tum.de

KURZFASSUNG

Datenpannen haben sich im letzten Jahrzehnt zunehmend zu einem schwerwiegenden Problem entwickelt. Obwohl der Diebstahl von Daten nicht immer nur aus finanziellen Gründen erfolgt, entstehen meist hohe Kosten für die Geschädigten. Diese Arbeit ordnet den Begriff Datenpanne ein, stellt aktuelle Zahlen vor und beschäftigt sich mit der Frage, wie sich der Trend in den letzten Jahren entwickelt hat. Dabei werden zwei verschiedene Datensammlungen genauer untersucht und Gemeinsamkeiten sowie vor allem Unterschiede vorgestellt. Insgesamt wird die Problematik der Vergleichbarkeit thematisiert. Zusätzlich wird der Umgang mit Datenpannen in der Realität anhand zweier bekannter Vorfälle gezeigt und daraus resultierende Herausforderungen, sowie die rechtliche Situation in verschiedenen Ländern und entstehende Kosten für Beteiligte analysiert.

Schlüsselworte

Datenpannen, LinkedIn Hack, Sony Pictures Entertainment Hack, Privacy Rights Clearinghouse, World's Biggest Data Breaches

1. EINLEITUNG

Seit 2005 werden Datenpannen von der Privacy Rights Clearinghouse Organisation in einer der größten Sammlungen an Datenlecks erfasst und analysiert. Bei der Betrachtung der Entwicklung der Datenmenge, die gehackt wurde bzw. anderweitig ungewollt abhanden kam, fällt auf, dass bis 2012 die Anzahl der Datenpannen gestiegen ist [8]. Allein im Jahr 2008 wurden so viele Datensätze gestohlen, wie in den vier vorhergehenden Jahren zusammen [30]. Dabei ist nicht jeder Verlust an Daten auf einen Hack zurückzuführen. Tatsächlich geht aus dem Datensatz der in [1] erfassten Datenpannen hervor, dass jeder zehnte Eintrag durch versehentliche Veröffentlichung entstand. Aus diesem Grund ist besonders für Unternehmen, die große Mengen an sensiblen Daten verwalten, interessant, welche Arten von Datenpannen auftreten können, welche Kosten entstehen, ob Trends in diesem Bereich feststellbar und welche Vorhersagen daraus ableitbar sind. Hat ein Unternehmen ein Datenleck entdeckt, so sind abhängig vom Standort außerdem gesetzliche Richtlinien bezüglich der Inkenntnissetzung der Betroffenen einzuhalten. In dieser Ausarbeitung soll deshalb zunächst ein Überblick rund um die Grundlagen von Datenpannen gegeben, sowie aktuelle Zahlen, Fakten und Trends veranschaulicht werden. Anschließend werden zwei konkrete Fallbeispiele im Detail betrachtet. In Kapitel 4 werden die verschiedenen Herausforderungen und Probleme insbesondere durch die entstandenen Kosten für Unternehmen, aber auch für die direkt Betroffenen diskutiert. Abschließend werden die Ergebnisse zusammengefasst und Erkenntnisse abgeleitet.

2. GRUNDLAGEN

Mit der steigenden Zahl an Datenpannen ist es aus analytischer Sicht von großer Bedeutung, sämtliche gemeldete Fälle sinnvoll zu erfassen und einzuordnen. Dabei gilt es in erster Linie zu überprüfen, ob es sich definitionsgemäß überhaupt um eine Datenpanne handelt und welcher Kategorie diese zuzuweisen ist. Hierfür werden in diesem Kapitel zunächst verschiedene Definitionen vorgestellt. Anschließend werden unterschiedliche Möglichkeiten zur Klassifizierung betrachtet. In einer kurzen Übersicht werden konkrete Zahlen und ein möglicherweise erkennbarer Trend veranschaulicht.

2.1 Definitionen

In der Literatur finden sich viele verschiedene Formulierungen, die den Begriff „Data Breach" beschreiben. In [28] wird z.B. folgende Definition festgelegt:

> „The term "data breach" means the loss, theft, or other unauthorized access, other than those incidental to the scope of employment, to data containing sensitive personal information, in electronic or printed form, that results in the potential compromise of the confidentiality or integrity of the data."

Diese Beschreibung bezieht einen (nicht unbedingt absichtlichen) Verlust explizit mit ein. Damit zählen beispielsweise auch verloren gegangene Datenträger als Datenpanne. Im Gegensatz hierzu wird in [9] folgende Definition verwendet:

> „A data breach is an incident where information is stolen or taken from a system without the knowledge or authorization of the system's owner."

Dabei beschränkt sich die Formulierung auf den aktiven Eingriff durch Stehlen oder anderweitiges Entwenden, sodass z.B. versehentlich abhanden gekommene Laptops nicht den Forderungen der Definition genügen. Für weitere Definitionen sei auf die Internationale Organisation für Normung (ISO)[16] und im weiteren Sinne auch auf das deutsche Bundesdatenschutzgesetz (BDSG)[6] verwiesen. Da in

doi: 10.2313/NET-2016-09-1_09

den meisten Sammlungen an Datenpannen durchaus zwischen absichtlichem Entwenden und versehentlichem Verlust unterschieden wird, wird in dieser Ausarbeitung die in [28] formulierte Begriffserklärung als Grundlage für die kommenden Kapitel verwendet.

2.2 Kategorien

Im Jahr 2013 wurde bekannt, dass Mitglieder der Social Media Plattform Facebook aufgrund eines Fehlers im System mit Hilfe eines Buttons ihre persönlichen Profilinformationen herunterladen konnten, die fälschlicherweise auch private Daten wie Telefonnummern und E-Mail Adressen von Freunden enthielten. Insgesamt waren rund 6 Millionen Datensätze betroffen[11]. Im selben Jahr wurden bis zu 2 Millionen Datensätze durch einen internen Mitarbeiter bei Vodafone entwendet[23]. Beide Fälle haben gemein, dass sie unter den Begriff „Datenpanne“ fallen. Sie unterscheiden sich jedoch in Art und zu Grunde liegendem Ablauf. Wie bereits in 2.1 angedeutet und auch in [26] zu finden, lassen sich diese in zwei große Kategorien einordnen: bösartig (malicious) und fahrlässig (negligent). Diese grobe Einteilung lässt sich wiederum, wie in Tabelle 1 aufgelistet, angelehnt an die in [1] verwendete Untergliederung weiter verfeinern. Während die-

Tabelle 1: Verfeinerung der Kategorien

Bösartig
Hacks
Insider
Diebstahl
Fahrlässig
Versehentlich veröffentlicht
Konfigurationsfehler
Verlust
Schwache Sicherheit

se Unterteilung den Fokus auf die Intention der am Datenleck Beteiligten legt, wird in [30] nach Art der Herkunft kategorisiert. Dabei zählen Angriffe von Hackern, kriminellen Organisationen und auch höhere Gewalt wie das Wetter zur Kategorie *Extern*, Datenpannen von Mitarbeitern innerhalb eines Unternehmens wie z.B. Administratoren mit hohen Zugriffsrechten zu *Intern* und von Partnerunternehmen ausgehende Attacken zu *Partner*. Eine weitere Möglichkeit besteht darin in *Softwarebasiert* (Hacks und Malware), *Hardwarebasiert* (Diebstahl und Verlust) und *menschliche Unachtsamkeit* (Fehler und versehentliche Weiterleitung)[27] zu unterteilen. Generell kann eine Datenpanne auch eine Verkettung mehrerer Punkte aus verschiedenen Kategorien umfassen. Dies wird beim Betrachten des auf das LinkedIn Profil des Facebookinhabers Mark Zuckerberg ausgeführten Hacks in Abschnitt 3.2 deutlich.

Jede der genannten Methoden zur Einteilung der Arten von Data Breaches hat Vor- und Nachteile. Für die schnelle und einfache Erfassung aller gemeldeten Fälle ist eine möglichst grobe Kategorisierung besser geeignet, als eine eher detaillierte Aufschlüsselung, insbesondere dann, wenn Begriffe wie Hacks noch weiter in einzelne Unterscheidungen wie Malware, Brute Force und SQL Injection zerlegt werden. Jedoch könnte eine nur sehr grobe Einteilung spätere Analysen erschweren, wenn beispielsweise gezielt Informationen über die Anzahl an Brute Force Angriffen gewünscht sind oder möglichst genaue Prognosen für die Zukunft abgeleitet werden sollen.

2.3 Datentypen

Um entscheiden zu können, wie schwerwiegend eine Datenpanne ist, muss bei der Erfassung der Daten anhand ihres Datentyps unterschieden werden. In der von [1] erstellten Sammlung wird hierbei jeder Datensatz mit Hilfe eines Sensibilitätswertes gekennzeichnet. E-Mail Adressen und Online Informationen wie beispielsweise Benutzernamen (ohne Passwort) werden als eher harmlos eingestuft (Wert 1), während Kreditkartennummern (Wert 3000) und E-Mail Passwörter (Wert 4000) als besonders sensible Daten gesehen werden. Laut dem California Data Breach Record 2016 sind Sozialnummern mit einer Gesamtheit von 24 Millionen Datensätzen die am häufigsten durch Datenpannen betroffene persönliche Information [27]. 39% der betrachteten Fälle enthalten Kreditkarteninformationen, 19% medizinische Daten [27].

2.4 Allgemeiner Ablauf

Je nach Art der Datenpanne lässt sich ein Muster bezüglich des Ablaufs erkennen. Bei Betrachtung der Angriffe der Kategorie *Bösartig* aus Tabelle 1 lassen sich diese mit drei aufeinanderfolgenden Schritten beschreiben [9]. Generell unterscheidet sich die Länge der einzelnen Schritte stark [30].

Schritt 1: Recherche
Der Angreifer sucht nach Möglichkeiten in das Zielsystem einzudringen. Die dafür eingesetzte Vorbereitungszeit bewegt sich hauptsächlich zwischen mehreren Stunden bis hin zu Monaten [30].

Schritt 2: Angriff
Für gefundene Schwachstellen im Bereich der Infrastruktur werden gezielt Angriffe in Form von beispielsweise SQL Injection oder Session Hijacking eingesetzt. Weitere Varianten sind Täuschungsversuche mit Hilfe von Phishing oder Spam. Für detaillierte Erklärungen dieser Angriffe sei auf [25] verwiesen. In 77% der Fälle erfolgt der Zugriff dabei innerhalb von wenigen Tagen, knapp ein drittel erfordert nicht einmal eine Stunde [30].

Schritt 3: Datentransfer
Sobald der Zugriff in Schritt zwei erlangt wurde, erfolgt der eigentliche Diebstahl der Daten, indem der Angreifer die Daten aus dem Zielsystem heraus transferiert. Die Zeitspanne hierfür erstreckt sich oft bis zu dem Zeitpunkt, an dem der Angegriffene die Attacke bemerkt und unterbindet. Bei knapp 50% der Datenpannen läuft diese Phase sogar über mehre Monate, wobei zwischen dem Entdecken des Problems und Schließen der Lücke in 42% der Fälle nochmals Wochen vergehen [30].

Kommen Daten direkt durch intern arbeitendes Personal abhanden, so erfolgen die Schritte eins und zwei implizit, da meist der Zugriff bereits besteht. Diese Art eines Angriffs ist demnach auch für Unternehmen oftmals am Schwierigsten zu verhindern. Für den Ablauf von Datenpannen der

Kategorie *Fahrlässig* ist jedoch oft kein eindeutiges Muster erkennbar. Generell gilt, dass für solch einen Vorgang zunächst eine Unachtsamkeit benötigt wird, die anschließend von einer kriminellen Instanz ausgenutzt wird. Zur Vermeidung der im ersten Schritt erwähnten Fahrlässigkeit sind allerdings Tipps wie ausführliches Testen, ständige Wartung und regelmäßige Aktualisierungen der Sicherheitsstandards der Systeme hilfreich. Weitere Vorschläge finden sich auch in [9]. Insbesondere ist jedoch das Wählen sicherer und einzigartiger Passwörter, die sich für sämtliche Zugänge unterscheiden sollten, ein wichtiger Aspekt, der auch in 3.2 eine große Rolle spielt.

2.5 Aktuelle Zahlen und Trends

Nun stellt sich auch für viele Unternehmen die Frage, ob die Häufigkeit und Größe von Datenpannen tatsächlich immer weiter steigt und wie „schlimm“ die Entwicklung wirklich ist. In diesem Kapitel wird einerseits die Datengrundlage der Analyse vorgestellt und aus dieser anschließend Trends abgeleitet.

2.5.1 Datensammlungen im Detail

Im Folgenden werden zwei Datensammlungen genauer beschrieben, die bereits in Kapitel 1 erwähnt wurden und besonders für die weitere Betrachtung in den Abschnitten 2.5.2 und 2.5.3 verwendet werden. Beide Datensammlungen haben gemein, dass die zugrunde liegenden Datensätze frei zugänglich und online zur Verfügung stehen und damit besser analysiert und verglichen werden können. Im Gegensatz dazu wird in [30] auf eine interne Datensammlung verwiesen, die nicht für eigene Untersuchungen zugänglich ist.

Privacy Rights Clearinghouse Organisation
Im Jahr 1992 gründete Beth Givens die kalifornische non-profit Organisation Privacy Rights Clearinghouse (PRC) mit dem Ziel Konsumenten über Datenschutz zu informieren und deren Privatsphäre zu verteidigen [8]. Seit 2005 wird eine über deren Webseite öffentlich zugängliche Sammlung von Datenpannen dokumentiert und gepflegt. Insgesamt umfasst die Datenbank 4860 Einträge mit knapp 90 Millionen betroffenen Datensätzen [8]. Abbildung 1 zeigt die Anteile der verschiedenen Kategorien an der Gesamtheit der Datenpannen auf Grundlage der PRC Datensammlung. Dabei wird deutlich, dass Hacks und Malware mit Abstand die häufigste Ursache von Datendiebstählen sind [8].

World’s Biggest Data Breaches (WBDB)
Der Autor, Journalist und Designer David McCandless hat sich auf die Visualisierung von großen Datenmengen spezialisiert und verschiedene Projekte mit Zugriff auf die zugrunde liegende Datenmenge online zur Verfügung gestellt. Dabei werden insbesondere Datenpannen mit über 30.000 betroffenen Datensätzen erfasst. Obwohl lediglich 218 Einträge vorliegen, ergibt sich eine Gesamtmenge von über 2,6 Milliarden betroffenen Datensätzen. Die Sammlung beschreibt daher Pannen, die als besonders schwerwiegend erachtet werden können, da eine große Menge an Daten pro Eintrag betroffen war [1].

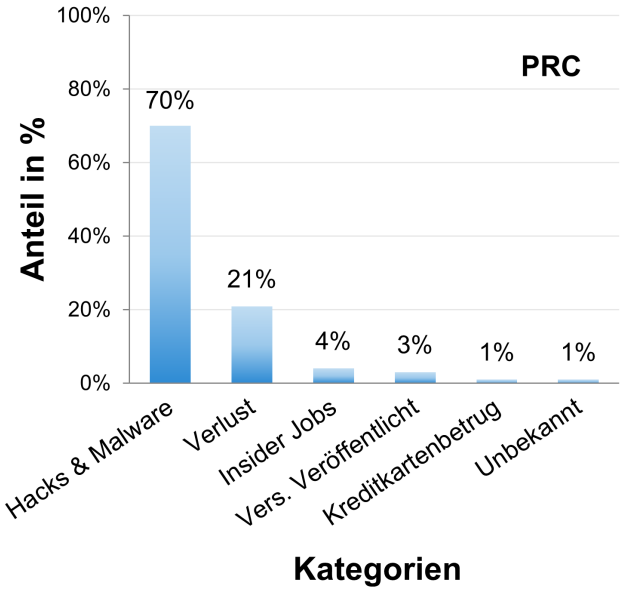

Abbildung 1: Anteil der einzelnen Kategorien an der Gesamtanzahl an Datenpannen der PRC Datensammlung [8]

2.5.2 Vergleich der Datenbanken

Im direkten Vergleich erscheinen die genannten Mengen der gestohlenen Daten zunächst widersprüchlich. Beim genaueren Untersuchen fällt jedoch auf, dass die PRC Organisation nur dann Werteangaben über die Höhe der betroffenen Datensätze einpflegt, sofern diese tatsächlich bestätigt wurden und es sich bei dem Verlust um sensible Daten handelt [8]. Für den in Abschnitt 3.1 betrachteten Fall gibt es beispielsweise keine offiziell bestätigten Quellen mit konkreten Mengenangaben. Die Angreifer jedoch behaupten über 100 Terabyte an Daten entwendet zu haben [4]. In der Verteilung der Anzahl der gestohlenen Datensätze pro eingetragener Datenpanne lassen sich ebenfalls Unterschiede erkennen, wie Abbildung 2 veranschaulicht. Im oberen Diagramm ist deutlich erkennbar, dass nur eine geringe Anzahl an Datenpannen eine wirklich große Menge an Datensätzen enthält. Bei Betrachtung der Datenbank von WBDB im unteren Diagramm fällt hingegen auf, dass viele der Datenpannen zum Verlust großer Datenmengen führten. Dies lässt sich durch die verschiedenen zugrunde liegenden Modelle beider Sammlungen erklären, da, wie im Namen des World’s Biggest Data Breaches Datensatzes enthalten, dort nur die größten Fälle erfasst und auch nicht eindeutig bestätigte Vorfälle aufgenommen werden.

2.5.3 Trends

In [26] wurde die Entwicklung der Häufigkeit und Größe an Datenpannen mit Hilfe der Daten des PRC untersucht, um passende Verteilungsfunktionen zu finden, die den Verlauf in den letzten Jahren möglichst genau beschreiben. Damit wurden auch Ansätze für verschiedene Vorhersagen vorgestellt. Die dadurch erlangten Erkenntnisse sind allerdings stark vom gegebenen Datensatz abhängig und könnten aufgrund der in 2.5.2 genannten Unterschiede völlig anders für die von WBDB zur Verfügung gestellten Datensammlung

doi: 10.2313/NET-2016-09-1_09

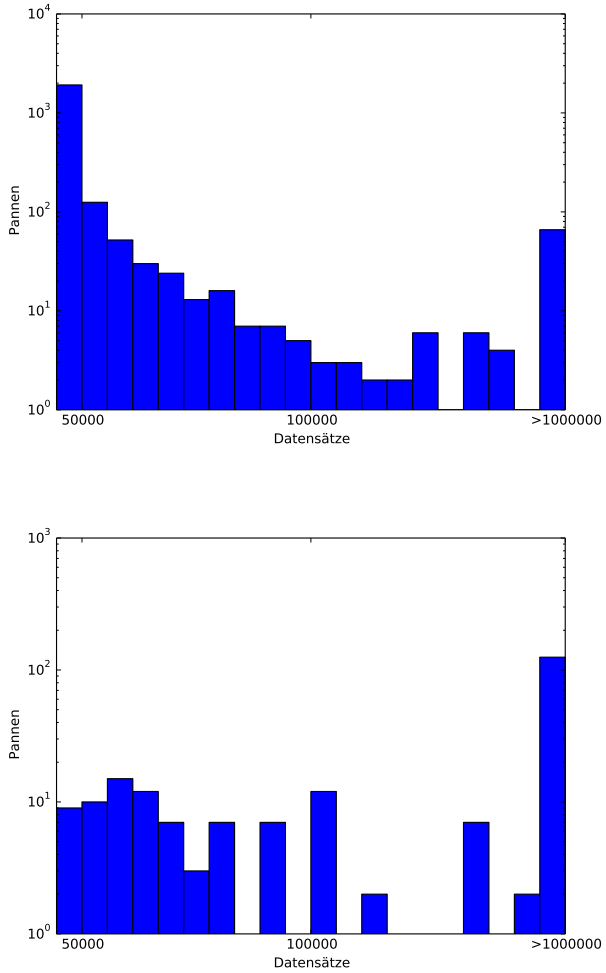

Abbildung 2: Verteilung der betroffenen Datensätze pro eingetragene Datenpanne. Zugrundeliegende Daten des PRC (oben) [8] und Datenbank von WBDB (unten) [1].

aussehen. Insofern lässt sich schwer festlegen, ob die Vorhersagen auch wirklich der Realität entsprechen. Die Betrachtung eines zweiten Datensatzes könnte die Genauigkeit jedoch verbessern. Zusätzlich werden Aussagen über die Entwicklung von Datenpannen auf Basis der PRC Datensammlung abgeleitet. Dabei wird insbesondere in der Menge an *bösartigen* Angriffen ein Rückgang und zeitgleich konstante Häufigkeit über einen Zeitraum von zehn Jahren ermittelt [26], während in anderen Quellen auf einen ansteigenden Trend der Menge an Datenpannen hingedeutet wird [15]. Oft trifft dies nur auf bestimmte Bereiche wie beispielsweise den Gesundheitssektor zu, während in anderen Gebieten die Zahl wiederum rückläufig ist. So zeigt auch der aktuelle California Data Breach Bericht von 2016, dass von 2012 bis 2015 die Menge an durch Hacks verursachten Datenpannen von 45% auf 58% stieg, jedoch gleichzeitig die Anzahl der auf Diebstahl oder Verlust zurückzuführenden Data Breaches um 10% sank. Abbildung 3 zeigt die Anzahl an Datenpannen über die letzten zwölf Jahre. Auch wenn es den Eindruck erweckt, als gäbe es ab 2012 einen eher gleichbleibenden bis sogar rückläufigen Trend bei der Menge an erfassten Datenpannen, sind für eine derartige Aussage zu wenig Datenpunkte vorhanden, um wirklich von einem „Trend“sprechen zu können. Insbesondere können durch die in Abschnitt 2.4 aufgezeigten Zeiten über die Länge der Dauer bis eine Lücke überhaupt erst entdeckt wird noch nachträgliche Eintragungen in der Datensammlung erfolgen, die bereits vergangenen Jahren zuzuordnen sind und die Entwicklungskurve aus Abbildung 3 nochmals verändern könnten. Andererseits wird beispielsweise in [2] erwähnt, dass in Deutschland mit Einführung der Informationspflicht Unternehmen einen erhöhten Fokus auf Sicherheitsaspekte legen, um dem steigenden Druck durch die mediale Aufmerksamkeit für Sicherheitslücken entgegen zu wirken. Derartige Gesetzesänderungen können somit Einfluss auf die Anzahl der erfassten Datenpannen nehmen.

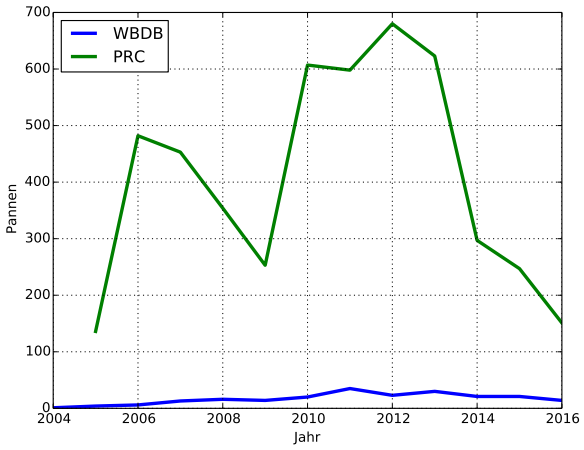

Abbildung 3: Entwicklung der Anzahl an Datenpannen von 2004 bis 2016 basierend auf den Datensätzen PRC [8] und WBDB [1].

3. FALLBEISPIELE

Nachdem in Kapitel 2 die Grundlagen und Begrifflichkeiten betrachtet wurden, sollen nun einige Fallbeispiele im Detail beleuchtet werden. Dabei wird auch deutlich, dass die Gründe und Folgen in manchen Fällen über typische Datenverluste hinaus gehen und sogar politische Bedeutung finden.

3.1 Sony Pictures Entertainment Hack

Der Sony Pictures Entertainment Hack erhielt im Jahr 2014 große mediale Aufmerksamkeit. Dabei sorgten nicht nur die Menge an gestohlenen Daten für große Diskussionen, sondern auch die Herkunft der mutmaßlichen Drahtzieher und deren angeblich politische Motivation. Dieser Vorfall veranschaulicht daher insbesondere, welche Konsequenzen Datenpannen auf internationaler Ebene nach sich ziehen können.

3.1.1 Beschreibung

Die zeitliche Entwicklung des Skandals um den Sony Hack im November 2014 erstreckte sich über eine Zeitspanne von etwa einem Monat. Eine genaue Auflistung der Ergebnisse

doi: 10.2313/NET-2016-09-1_09

in chronologischer Reihenfolge wurde in [22] erarbeitet. Obwohl bekannt ist, dass die Datenpanne auf eine Sicherheitslücke zurückzuführen ist, über die Malware in das System von Sony Entertainment Pictures eingebracht wurde, konnte bis heute nicht abschließend geklärt werden, wie viele Daten genau gestohlen wurden. Die Berichterstattung begann damit, dass Mitarbeiter der *Sony Pictures Entertainment Inc.* aufgrund einer Sperre ihrer Arbeitsrechner durch ein Bild mit dem Text „Hacked By #GOP“ an der Aufnahme ihrer Arbeit gehindert wurden. Die zuständige IT-Abteilung stellte daraufhin sogar fest, dass alle der Infrastruktur angehörenden Rechner, gesperrt waren und ein Teil der im Unternehmensnetzwerk befindlichen Daten zerstört wurde. Es folgte eine Kontaktaufnahme der Gruppe *Guardians of Peace (#GOP)*, welche angab, die genannte Sperre installiert zu haben, und außerdem behauptete, mehrere 100 Terabyte an Daten aus dem Firmennetzwerk entwendet zu haben. Als Beweis wurden durch GOP fünf von Sony Pictures Entertainment produzierte Filme online verbreitet, die noch nicht veröffentlicht und daher nur unternehmensintern zugänglich waren. Die kurze Zeit später gestellten Forderungen der Gruppe richteten sich einerseits an Mitarbeiter des Konzerns, die sich öffentlich von Sony Pictures Entertainment distanzieren sollten, andererseits an Sony selbst, mit dem Verlangen, die Veröffentlichung des Films *The Interview* zu unterlassen. Zur Erhöhung des Drucks wurden laufend weitere interne Daten des Konzerns durch die Hackergruppe veröffentlicht. Der genannte Film stellt dabei ein kritisches Werk gegenüber dem nordkoreanischen Staatsoberhaupt Kim Jong-un dar und wurde in Nordkorea derart negativ beurteilt, dass dort die Veröffentlichung des Bildmaterials sogar als „Kriegshandlung“ und „unverhohlene Unterstützung von Terrorismus“ eingestuft wurde [19]. Es erhärtete sich somit der Verdacht, dass die Regierung Nordkoreas bei dem Angriff auf Sony Pictures Entertainment involviert war. Dies führte sogar dazu, dass das FBI nach einer Untersuchung des Falls offiziell bestätigte, dass die Regierung Nordkoreas den Angriff durchgeführt hätte, was von dieser jedoch sofort dementiert wurde. Bis heute konnte eine Beteiligung Nordkoreas nicht zweifelsfrei nachgewiesen werden. Aufgrund des enormen Drucks, der durch die Gruppe #GOP ausgeübt wurde, entschieden sich die Verantwortlichen von Sony zwischenzeitlich gegen eine Veröffentlichung des Films, annullierten diese Entscheidung jedoch kurze Zeit später und willigten einer Ausstrahlung des Werkes in Kinos ein.

3.1.2 Vorgehensweise

Zur Vorgehensweise der Angreifer wurden nie Informationen explizit im Bezug zum Sony Pictures Entertainment Hack veröffentlicht. Jedoch veröffentlichte das *US-CERT*, eine US-Organisation des Department of Homeland Security, eine Stellungnahme, in der beschrieben wird, dass eine Malware namens *SMB Worm Tool* verwendet wurde, um Attacken auf ein „großes Unterhaltungsunternehmen“ auszuführen. Die Vermutung liegt nahe, dass es sich dabei um den Sony Pictures Entertainment Hack handelt, da die beschriebene Schadsoftware sämtliche Funktionalität aufweist, um den auf die Infrastruktur des Unternehmens ausgeführten Angriff durchzuführen. Eine genaue Beschreibung des Wurms wurde in [14] veröffentlicht.

Die Software wurde auf unbekannte Weise in die Infrastruktur des Konzerns eingebracht und besitzt die Eigenschaft, sich über das *Server Message Block (SMB)* Protokoll im gesamten Netzwerk zu verbreiten. Sie besteht aus mehreren Modulen, die einerseits Funktionalität zur Durchführung von Brute Force Angriffen auf SMB-Freigaben zur Verfügung stellen, um eine Verbreitung im Netzwerk zu ermöglichen. Andererseits werden Funktionen zur Verfügung gestellt, um mit einer *command and control (C2)* Infrastruktur zu kommunizieren, die einen Informationsaustausch und die Steuerung der Malware durch die Hacker von außerhalb ermöglicht. Zusätzlich installierte die Software Module, die Daten im infizierten Netzwerk unwiederbringlich löschen konnten bzw. die gesamte Festplatte eines Hostsystems durch die Zerstörung des Master Boot Records (MBR) ermöglichte [5].

Der gesamte Ablauf des Angriffs erfolgte also in drei Schritten. Zuerst wurde die Malware auf ungeklärte Weise in das Netzwerk der Firma Sony geschleust. Es wird vermutet, dass dafür ein Phishing Angriff durchgeführt wurde [21]. Daraufhin wurde die Schadsoftware über SMB-Freigaben im Netzwerk verteilt und Daten abgegriffen. Nach Angaben der Gruppe #GOP dauerte die Ausführung dieses Schritts mindestens ein Jahr [22]. Schließlich wurden Daten im Netzwerk gelöscht und die infizierten Rechner unbrauchbar gemacht.

3.1.3 Folgen

Vor allem aufgrund des hohen politischen Interesses am Sony Pictures Entertainment Hack hebt sich der Fall klar von anderen Datenpannen ab. Der Diebstahl von Daten und die Verwendung dieser für Erpressungsversuche führte zwischenzeitlich sogar dazu, dass die Regierung der Vereinigten Staaten von Amerika eine Verfügung des Präsidenten veröffentlichte, die zusätzliche Sanktionen gegen Nordkorea ankündigte und dabei explizit Bezug auf die Attacke gegen Sony nahm [10].

3.2 LinkedIn und Mark Zuckerberg

Anfang Juni 2016 erregte ein Hackerangriff auf die Accounts verschiedener prominenter Persönlichkeiten auf sozialen Plattformen Aufsehen. Neben Berühmtheiten wie Rolling-Stones-Musiker Keith Richards und Reality-TV-Star Kylie Jenner war auch Mark Zuckerberg, der Gründer und Vorstandsvorsitzende der größten Social Media Plattform *Facebook*, betroffen [20]. Der Fall brachte eine besondere Brisanz mit sich, da es der breiten Masse möglich war, den Verlauf des Angriffs über die betroffenen sozialen Medien zeitnah zu verfolgen.

3.2.1 Beschreibung

Obwohl bei den Vorfällen im Jahr 2016 der Hergang zuerst nicht klar war, stellte sich später heraus, dass die zugrundeliegende Ursache schon auf das Jahr 2012 zurück ging. Damals wurde seitens LinkedIn, einem sozialen Netzwerk, das überwiegend auf berufliche Kontakte ausgelegt ist, bekannt gegeben, dass ein Einbruch in die Datenbank stattgefunden hatte und dabei die Passwörter „einiger Mitglieder“ entwendet wurden [17]. Vier Jahre später wurde klar, dass es sich bei „einigen“ Mitgliedern um etwa 6 Millionen Nutzer handelte, deren Zugangsdaten gestohlen wurden. Diese wurden verwendet, um gefälschte Informationen über die Accounts öffentlicher Persönlichkeiten zu verbreiten.

3.2.2 Vorgehensweise

Bis heute ist nicht abschließend geklärt wie sich Hacker Zugriff auf die Nutzerdatenbank bei LinkedIn verschaffen konnten. Jedoch wurden Informationen veröffentlicht, die belegen, dass die Nutzerpasswörter als ungesalzener SHA1-Hash vorlagen [13]. Grundsätzlich ist die Vorgehensweise richtig, Passwörter nicht im Klartext zu speichern, sondern durch eine Einwegfunktion (Hash) unkenntlich zu machen. Bei der Umsetzung wurde jedoch in diesem Fall auf einen Salt-Wert verzichtet, einer zufälligen Zeichenkette, die an jede Eingabe der Hashfunktion angehängt wird, um die Verwendung von vorberechneten Tabellen zur Wiederherstellung des Klartextes aus dem berechneten Hashwert zu erschweren [29]. Aus den mutmaßlich 6 Millionen entwendeten Passworthashes konnten daher innerhalb weniger Tage die Passwörter im Klartext ermittelt werden. Auf der Plattform LinkedIn selbst wurden zwar alle Passwörter zeitnah zurückgesetzt, doch wurde die Gefahr unterschätzt, dass viele Nutzer ein und das selbe Passwort auf vielen Social Media Plattformen verwenden. Es war den Angreifern damit möglich, sich mit den von LinkedIn entwendeten Passwörtern Zugriff auf Accounts in anderen Portalen zu verschaffen. Auch Mark Zuckerberg verwendete auf LinkedIn, Twitter und dem Photodienst Pinterest das selbe Passwort, das unbestätigten Berichten zufolge im Klartext „dadada“ lautete [18]. Die Hacker hatten damit die Möglichkeit, Falschinformationen über die gehackten Accounts zu verbreiten. Außerdem wurde Zuckerbergs Pinterest Auftritt derart geändert, dass die Profilbeschreibung den Text „gehackt vom OurMine Team“ enthielt [12].

3.2.3 Folgen

Aufgrund der Tatsache, dass die Identität der Gruppe, die unter dem Namen *OurMine Team* operiert, ungeklärt ist, sind in Sachen Strafverfolgung keine Ergebnisse bekannt. Die betroffenen sozialen Netzwerke reagierten mit einer Sperrung der Accounts, die mit der Hackgergruppe in Verbindung gebracht wurden. Ein Aufruf des Twitter-Profils *_OurMine_* wird automatisch auf die Website `https://twitter.com/account/suspended` umgeleitet, die stets bei dem Versuch aufgerufen wird, das Profil eines bei Twitter gesperrten Accounts zu besuchen.

4. HERAUSFORDERUNGEN

Wie in Abschnitt 2.4 bereits erwähnt, kann sich der Zeitraum vom Beginn eines Angriffs bis hin zur Entdeckung und Eindämmung über mehrere Monate, oder sogar Jahre erstrecken, was für alle Beteiligten besondere Herausforderungen birgt. Im Falle einer Datenpanne ist nicht nur wichtig diese möglichst schnell zu entdecken und weitere Zugriffe zu unterbinden, sondern auch Betroffene zeitnah zu informieren und entstandene Schäden zu begrenzen. Dies gelingt beteiligten Unternehmen nicht immer, was Folgeangriffe, wie im Fallbeispiel in Abschnitt 3.2 gezeigt, erst möglich macht und auch Jahre nach dem Angriff noch zu Konsequenzen führen kann. Derartige Skandale können außerdem hohe finanzielle Einbußen nach sich ziehen. Im Folgenden werden daher die Verantwortlichkeiten der Firmen erläutert und anschließend näher auf Art und Höhe der entstehenden Kosten eingegangen.

4.1 Rechtliche Meldepflicht

Da kein international geltendes Recht für Datenpannen existiert, hängt der korrekte Umgang mit diesen von der Rechtsprechung des jeweiligen Standortes ab. Dies erschwert die Beurteilung der Fälle zusätzlich, da keine einheitliche Rechtsauffassung zugrunde gelegt werden kann. In den Vereinigten Staaten von Amerika variieren die rechtlichen Vorgaben sogar von Staat zu Staat, wobei 47 der 51 Staaten eine Informationspflicht bei Verlust von persönlichen Daten gegenüber Betroffenen vorschreiben. 15 Staaten haben dabei Regularien bezüglich der Form der Informationsmitteilung festgelegt, wie in [24] in tabellarischer Form dargestellt. In Deutschland sind die Rechte und Pflichten im Bundesdatenschutzgesetz geregelt. Dort ist eine Informationspflicht bei Auftreten von Datenpannen festgelegt. Seit der Aufnahme der neuen Vorschrift in §42a stieg das Interesse der Firmen, sich stärker im Bereich IT-Sicherheit zu positionieren[2]. Dabei ist auch auf die Form der Meldung zu achten, die vorgibt, dass ein Unternehmen eine Datenpanne in „mindestens eine halbe Seite umfassen, [und] in mindestens zwei bundesweit erscheinenden Tageszeitungen“ [6] publiziert werden muss.

4.2 Kosten

Wird gegen die im vorhergehenden Abschnitt 4.1 genannten Pflichten verstoßen, so drohen hohe Bußgelder bis zu 300.000€[7]. Insbesondere leidet das für viele Firmen wichtige Image unter mangelhafter Aufklärung, was Kosten in nicht absehbarer Höhe verursachen kann. Neben einer genaueren Analyse der Kostenentwicklung werden in [3] die finanziellen Aufwendungen in vier Kategorien eingeteilt. Dabei wird zwischen *Erkennung und Eskalierung* (Ermittlung der Faktenlage und interne Kommunikation), *Benachrichtigung* (Publikation der Panne), *ex-post Kosten* (Nachbereitung und Schließung der Lücken) und *Businessverlust* (Abgang von Kunden und Geschäftspartnern) unterschieden. Statistisch gesehen stiegen die finanziellen Aufwendungen jedes Jahr stetig an, während Businessverlust und ex-post Kosten den größten Teil ausmachen. Auch der durchschnittliche finanzielle Gesamtschaden pro Datenpanne stieg über die Jahre bis auf 3.79 Millionen US-Dollar (2015) [3].

5. ZUSAMMENFASSUNG

In dieser Arbeit wurden die Grundlagen zu Datenpannen vorgestellt und gängige Definitionen als Basis für weitere Analysen dargestellt. Dabei wurde auch auf die unterschiedlichen Interpretationen des Begriffs bei der Klassifikation bekannter Fälle eingegangen und zwei Datensammlungen analysiert, die auf verschiedene Weise die Pannen erfassen. Es zeigte sich, dass die resultierenden Datenbanken in hohem Maße abhängig von der Einordnung bezüglich unbestätigter Informationen und Vollständigkeit der zugrundeliegenden Datensätze sind. Mit dieser Erkenntnis wurde ein Trend herausgestellt, der einen Anstieg der Häufigkeit von Datenpannen bis 2012 beschreibt. Für die Bestimmung eines Trends der Folgejahre liegen jedoch genügend Daten vor. Es folgte die Aufarbeitung zweier brisanter Beispiele, die zeigten, dass die Entstehung von Pannen sowohl durch eine mangelhafte Sicherheitspolitik bei Unternehmen, als auch durch geringes Sicherheitsbewusstsein beim Kunden selbst massiv begünstigt wird. Abschließend wurden die für Firmen auftretenden Herausforderungen benannt und die geltende Gesetzgebung sowie die entstehenden Kosten untersucht. Es zeigte sich, dass die Einführung von Gesetzen zur Verbesserung

doi: 10.2313/NET-2016-09-1_09

der Unternehmenstransparenz zu einem erhöhten Fokus auf IT-Sicherheit in Firmen führt.

Durch die starken Unterschiede im Design von Datensammlungen lassen sich diese nur bedingt vergleichen. Eine einheitliche und international greifende Regelung zur Einordnung von Datensätzen, sowie global geltende Gesetze zum Thema Meldepflicht wären daher wünschenswert. Interessant wäre zudem eine Anwendung und Analyse des in [26] vorgestellten Modells auf verschiedenen Datensammlungen, um dessen universelle Einsetzbarkeit zu beurteilen.

6. LITERATUR

[1] http://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/. 2016-05-20.

[2] 2010 Annual Study: German Cost of Data Breach. http://www.symantec.com/content/de/de/about/downloads/press/2010_annual_study.pdf. 2016-06-19.

[3] 2015 Cost of Data Breach Study: Global Analysis. https://nhlearningsolutions.com/Portals/0/Documents/2015-Cost-of-Data-Breach-Study.PDF. 2016-06-19.

[4] A Look Through The Sony Pictures Data Hack: This Is As Bad As It Gets. https://www.buzzfeed.com/tomgara/sony-hack?utm_term=.rfBKoqaQv#.dukOe4VvQ. 2016-06-18.

[5] Alert (TA14-353A) Targeted Destructive Malware. https://www.us-cert.gov/ncas/alerts/TA14-353A. 2016-06-17.

[6] Bundesdatenschutzgesetz (BDSG) § 42a Informationspflicht bei unrechtmäßiger Kenntniserlangung von Daten. https://www.gesetze-im-internet.de/bdsg_1990/__42a.html. 2016-06-16.

[7] Bundesdatenschutzgesetz (BDSG) § 43a Informationspflicht bei unrechtmäßiger Kenntniserlangung von Daten. https://www.gesetze-im-internet.de/bdsg_1990/__43.html. 2016-06-19.

[8] Chronology of data breaches security breaches 2005 - present. http://www.privacyrights.org/data-breach. 2016-05-20.

[9] Definition - data breach. http://www.trendmicro.com/vinfo/us/security/definition/data-breach. 2016-05-20.

[10] Executive Order – Imposing Additional Sanctions with Respect to North Korea. https://www.whitehouse.gov/the-press-office/2015/01/02/executive-order-imposing-additional-sanctions-respect-north-korea. 2016-06-17.

[11] Facebook admits year-long data breach exposed 6 million users. http://www.reuters.com/article/net-us-facebook-security-idUSBRE95K18Y20130621. 2016-06-16.

[12] Hacker legen Mark Zuckerberg rein. http://www.rp-online.de/digitales/internet/cyper-attacke-hacker-nehmen-facebook-chef-mark-zuckerberg-ins-visier-aid-1.6027071. 2016-06-17.

[13] Hacker puts up 167 Million LinkedIn Passwords for Sale. http://thehackernews.com/2016/05/linkedin-account-hack.html. 2016-06-17.

[14] Hackers Used Sophisticated SMB Worm Tool to Attack Sony. http://www.securityweek.com/hackers-used-sophisticated-smb-worm-tool-attack-sony. 2016-06-17.

[15] Internet Security Threat Report 2014. http://www.symantec.com/content/en/us/enterprise/other_resources/b-istr_main_report_v19_21291018.en-us.pdf. 2016-06-19.

[16] ISO/IEC 27040. https://www.iso.org/obp/ui/#iso:std:iso-iec:27040:ed-1:v1:en. 2016-06-16.

[17] LinkedIn resets passwords on millions of accounts as new data-leak reports surface. http://venturebeat.com/2016/05/18/linkedin-resets-passwords-on-millions-of-accounts-as-new-data-leak-reports-surface/. 2016-06-17.

[18] Mark Zuckerberg's password was 'dadada'. What hope do the rest of us have? http://www.telegraph.co.uk/technology/2016/06/06/mark-zuckerbergs-password-was-dadada-what-hope-do-the-rest-of-us/. 2016-06-17.

[19] Nordkorea protestiert gegen Kinofilm über Kim Jong-un. http://www.faz.net/aktuell/politik/ausland/brief-an-un-generalsekretaer-nordkorea-protestiert-gegen-kinofilm-ueber-kim-jong-un-13038080.html. 2016-06-17.

[20] Passwortsünder Zuckerberg? http://www.golem.de/news/social-media-passwortsuender-zuckerberg-1606-121326.html. 2016-06-17.

[21] Sony Got Hacked Hard: What We Know and Don't Know So Far. https://www.wired.com/2014/12/sony-hack-what-we-know/. 2016-06-18.

[22] Sony Hack: A Timeline. http://deadline.com/2014/12/sony-hack-timeline-any-pascal-the-interview-north-korea-1201325501/. 2016-06-16.

[23] Vodafone Deutschland Ziel eines Angriffs. http://www.vodafone.de/privat/hilfe-support/kundeninformation.html. 2016-06-16.

[24] F. Bisogni. Data breaches and the dilemmas in notifying customers. In *The 14th Annual Workshop on the Economics of Information Security*, pages 23–25, 2015.

[25] W. A. S. Consortium et al. Threat classification. *Online at: http://www. webappsec. org/projects/threat/v1/WASC-TC-v1 0. pdf*, 2004.

[26] B. Edwards, S. Hofmeyr, and S. Forrest. Hype and heavy tails: A closer look at data breaches. 2015.

[27] K. D. Harris and A. General. California data breach report. 2016.

[28] O. of the Federal Register. *Code of Federal Regulations, Title 38, Pensions, Bonuses, and Veterans' Relief, Pt. 18-End, Revised as of July 1 2009*. U.S. Government Printing Office, 2009.

[29] D. Todorov. *Mechanics of User Identification and*

Authentication: Fundamentals of Identity Management. CRC Press, 2007.

[30] R. Verizon Business. Team, 2009 data breach investigations report, 2009.

doi: 10.2313/NET-2016-09-1_09