Die Wissenschaft des Ratens:

Analyse einer anonymisierten Sammlung von 70 Millionen Passwörtern

Konstantin Ruhmann

FH Wedel

Wintersemester 2014/2015

I Inhaltsverzeichnis

II Abbildungsverzeichnis

1. Einleitung

1.1. Zielsetzung

Passwörter sind in der Authentifizierung eine elementare Komponente. Besonders durch die starke Entwicklung des Internets nimmt die Bedeutung von Passwörtern stetig zu. Benutzer haben für Dienste im Internet meist ein Benutzernamen und ein Passwort um sich zu authentifizieren. Die Ausprägung eines Passworts ist grundsätzlich beliebig, doch sehen Sicherheitsforscher Passwörter kritisch.

In der Studie von Bonneau wird die Frage gestellt, wie sicher Passwörter gegen das Erraten sind. Dazu wurde in Zusammenarbeit mit dem Unternehmen Yahoo! eine Passwortsammlung erhoben und ausgewertet. In Summe wurden 70 Millionen Passwörter erhoben und jeweils mit 328 demographischen Informationen angereichert.

1.2. Abgrenzung

Die Ausarbeitung bezieht sich auf eine Publikation von Joseph Bonneau aus dem Jahr 2012. Der Autor dieser Ausarbeitung leitet das Thema zunächst mit einer allgemeinen Beschreibung des Begriffs Passwort und die allgemeine Nutzung von Passwörtern ein. Anschließend erfolgt eine historische Betrachtung von Studien die sich mit dem Cracking von Passwörtern bzw. Semantik von Passwörtern befasst. Die Passworterhebung bei Yahoo! erfolgt unter den Gesichtspunkten, die Sicherheit der Passwörter zu garantieren und die Passwörter mit demographischen Informationen anzureichern. In dem folgenden Kapitel werden die Kennzahlen für die Auswertung vorgestellt. Dabei wird sowohl auf fremde Kennzahlen, als auch auf Kennzahlen von Bonneau hingewiesen. Die Herleitung der mathematischen Formeln für die Kennzahlen ist nicht Teil der Ausarbeitung. Im Anschluss wird die Analyse der Passwortsammlung von Yahoo! vorgestellt.

2. Bedeutung der Passwortsicherheit

Ein Passwort besteht aus einer alphanumerischen Zeichenfolge für die es keine allgemeinen Beschränkungen gibt. Mögliche Beschränkungen sind beispielsweise die Passwortlänge, Verwendung von bestimmten Zeichen oder die Lebensdauer von einem Passwort. Beschränkungen werden von einem Administrator vorgegeben, aber der Benutzer kann auf Basis dieser Beschränkungen das Passwort für gewöhnlich frei wählen. Dies führt in der Praxis zu folgenden Herausforderungen.

Der Benutzer versucht Passwortkombinationen zu wählen, die er sich gut merken kann. Angreifer versuchen dann solche Passwörter mit Hilfe von Wörterbuchangriffen zu erraten. Wörterbuchangriffe basieren auf einer großen Sammlung von bekannten Passwörtern. Der Administrator sollte daher die Beschränkungen so wählen, dass die Anfälligkeit für Wörterbuchangriffe gemindert wird. Dies kann jedoch zu dem Effekt führen, dass die Passwörter für den Benutzer zu kompliziert werden und er sie vergisst bzw. das Passwort notiert und unter die Tastatur legt. Hier muss der Administrator eine geeignete Balance finden.

Typischerweise kann sich ein Anwender mit einem Passwort an einem Anwendungssystem anmelden. Aber auch eine einzelne Datei könnte mit einem Passwort geschützt werden. Das Passwort dient als Schlüssel zu einer geheimen Information. Dieser Prozess wird auch als Authentifizierung bezeichnet.

Authentifizierung beschreibt die Überprüfung ob eine Person diejenige ist, als die sie vorgibt zu sein. Dies erfolgt beispielweise bei der Anmeldung auf einer Webseite um seine Emails zu lesen. Hier authentifiziert sich ein Benutzer typsicherweise mit Hilfe einer Email-Adresse und einem Passwort. Dies wird auch als Authentifizierung durch Wissen bezeichnet. Das Passwort ist in diesem Fall das Wissen. Nach Abschluss der Authentifizierung ist der Zugriff auf die Webseite meist ohne weitere Einschränkungen möglich. Übertragen kann dies auch für weitere Anwendungssysteme gelten. Wer im Besitz des Wissens ist, kann sich am System legitim authentifizieren und die Informationen erhalten. Das Passwort ist damit für einen Angreifer eine Möglichkeit sich Zugang zu einem System zu schaffen.[1]

3. Historische Betrachtung

3.1. Cracking

Der Begriff Cracking beschreibt das Brechen von Passwörtern. Im Allgemeinen werden dazu Wörterbücher verwendet, die eine Sammlung von Passwörtern beinhalten. Die Größe der Sammlung und die Passwörter in der Sammlung haben Einfluss auf die Erfolgsrate die Passwörter zu brechen.

Bereits im Jahr 1979 wurde eine Analyse von 3000 Passwörtern durchgeführt. In diesem Fall wurde als Wörterbuch das Wörterbuch des Systems für die Analyse verwendet und die Analyse wurde auf 6-stellige Zeichenfolgen eingeschränkt. Im Ergebnis wurden 84% der Passwörter erraten. Dabei wurden auch die ersten statistischen Daten über Passwortlänge und Passwortzeichenfolgen in einer Passwortsammlung erhoben. Beispielweise waren von allen

[1] Vgl. Schmeh (2013), s.S. 405ff.

Passwörtern 71% kleiner oder gleich 6 Stellen und 14% der Passwörter basierten auf nicht alphanummerische Zeichen.[2]

Die allgemeine Aufmerksamkeit für das Ausprobieren von schwachen Passwörtern schaffte ein Wurm im Jahre 1988. Ein Wurm ist ein Schadecode, der sich durch Ausnutzen von Schwachstellen, eigenständig verbreiten kann. Eine Funktion des Wurms war das Erraten von einfachen Passwörtern mit Hilfe eines Wörterbuches, welches lediglich 350 Wörter umfasste. Die Anzahl der Elemente wurden jedoch durch Modifikationen der Wörter erweitert.[3] Diese Anpassungsregeln werden im Englischen als Mangling-Rules bezeichnet. Sie erweitern automatisch ein Wörterbuch, indem zum Beispiel Kleinbuchstaben durch Großbuchstaben ersetzt werden. Die Entwicklung von Mangling-Rules wurde durch unterstützende Anwendungen vorangetrieben, welche immer mehr Anpassungsregeln unterstützten. Diese Anwendungen wurden in späteren Studien auf immer größere Sammlungen von Passwörtern angewendet um dann die Rate der gecrackten Passwörter zu ermitteln.

In der Abbildung 1 werden verschiedene Studien aus verschiedenen Jahren dargestellt. Die Studien unterscheiden sich in der Größe der Wörterbücher und teils durch unterschiedliche Passwortsammlungen. Die Effizienz der einzelnen Studien variierte stark und der Großteil aller Studien erzielt 20-50% geknackte Passwörter. Bessere Erfolgsraten wurden durch Optimierung von Wörterbüchern erzielt. Eine erste Erkenntnis zeigte im Vergleich aller Studien jedoch, dass die Effizienz mit größeren Wörterbüchern abnimmt.

[2] Vgl. Bonneau (2012), s. S. 2.
[3] Vgl. Bonneau (2012), s. S. 2.

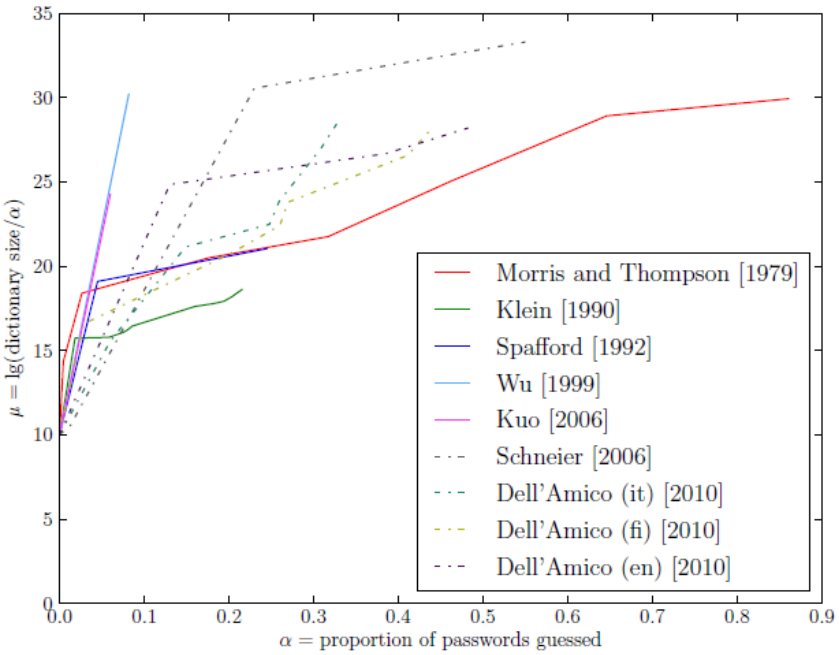

Abbildung 1: Vergleich Studien[4]

3.2. Semantik

Die Analyse der Semantik und Sprache von Passwörtern hat die Herausforderung, dass dazu Passwörter im Klartext benötigt werden. Zudem möchte man Beziehungen zwischen Passwort und Benutzer herstellen. Die Erhebung von Passwörtern kann beispielsweise durch Umfragen erfolgen, jedoch können bei diesem Vorgehen diese Ergebnisse von der Wirklichkeit abweichen, weil Benutzer durch die Fragestellung zu besseren Passwörtern animiert werden. Die Klartext-Passwörter der Benutzer ohne Erlaubnis zu nutzen widerspricht auf der anderen Seite der Privatsphäre der Benutzer.

Nach Erhebung der Passwörter können dann linguistische Analysen, wie beispielweise die Klassifizierung der Passwörter erfolgen. Solch eine Analyse hat bereits im Jahr 1989 stattgefunden und es wurden 6226 Passwörter klassifiziert, wobei einige lediglich als zufällige Zeichenfolgen klassifiziert wurden. Im Jahr 2006 wurde durch Cazier et al. festgestellt, dass schwer klassifizierbare Passwörter schwer zu erraten sind.[5]

Durch die Strukturierung von Passwörtern wurde es möglich Kennzahlen, wie die Länge, Anteil an Zahlen und Anteil an Sonderzeichen zu ermitteln. Die folgende Grafik zeigt die genannten Kennzahlen für die durchgeführten Analysen.[6]

[4] Vgl. Bonneau (2012), s. S. 2.
[5] Vgl. Bonneau (2012), s. S. 3.
[6] Vgl. Bonneau (2012), s. S. 3.

year	study	length	% digits	% special
1989	Riddle et al. [15]	4.4	3.5	—
1992	Spafford [5]	6.8	31.7	14.8
1999	Wu [12]	7.5	25.7	4.1
1999	Zviran and Haga [18]	5.7	19.2	0.7
2006	Cazier and Medlin [14]	7.4	35.0	1.3
2009	*RockYou leak* [19]	7.9	54.0	3.7

Abbildung 2: Vergleich Kennzahlen[7]

Die Kennzahlen der Studien haben zueinander große Abweichungen, sodass keine konkreten Aussagen, sondern lediglich Trends ermittelt werden können. Ein Trend ist beispielsweise eine Passwortlänge zwischen 6 und 8 Zeichen und die Benutzer nutzen gerne Buchstaben und meiden Zahlen bzw. Sonderzeichen. Einige Studien haben ebenfalls versucht Werte bezüglich zufälliger Passwörter zu ermitteln. Zufällige Passwörter sind Passwörter, die keinen Zusammenhang zwischen Passwort und Eigenschaften eines Benutzers haben. Hier wurden Werte zwischen 10 und 50 % erreicht.

Die Kennzahlen, wie Länge, Groß- bzw. Kleinschreibung und Sonderzeichen werden ebenfalls genutzt um die Passwortstärke zu ermitteln. Dieses Verfahren stammt von der Federal Information Processing Standard (FIPS), der Standardisierungsbehörde der Vereinigten Staaten, und wurde in der Richtlinie Electronic Authentication Guideline im Jahr 2006 beschrieben. Die Richtlinie ermittelt für Passwörter einen Wert, der die Charakteristiken eines Passwortes beschreibt.[8]

3.3. Probleme bei den gezeigten Ansätzen

Um ein wissenschaftliches Verständnis für Passwortsicherheit zu entwickeln, gibt es bei der Betrachtung der bisherigen Studien folgende Schwierigkeiten.

Zum einen ist die Vergleichbarkeit der Studien schwierig. Die Studien ermitteln unterschiedliche Kennzahlen, teilweise fehlen Kennzahlen und die Wörterbücher variieren in ihrer Größe. Beispielsweise haben einige Studien die Laufzeit von Programmen ermittelt, jedoch keine Informationen über die Wörterbuchgröße.

Ein weiteres Problem ist die Wiederholbarkeit der verschiedenen Studien. Die populäre Anwendung „John the Ripper“, welche von zahlreichen Studien für die Mangling-Rules verwendet wird, ist in den Jahren in zahlreichen Versionen erschienen, in denen jeweils unterschiedliche Konfigurationseinstellungen vorgenommen werden konnten. Andere Studien

[7] Vgl. Bonneau (2012), s. S. 4.
[8] Vgl. Bonneau (2012), s. S. 3.

haben eigene Anwendungen geschrieben, die der Öffentlichkeit nicht zur Verfügung stehen. Darüber hinaus haben die Studien speziell angepasste Wörterbücher verwendet, die eine Wiederholbarkeit kaum ermöglichen.

Neben den Anwendungen selbst haben die Wörterbücher elementare Auswirkungen auf die Ergebnisse. Speziell angepasste Wörterbücher können auf bestimmte Passwortsammlung deutlich bessere Ergebnisse erzielen. Bezüglich einer Passwortsammlung existieren auf diese Weise mehrdeutige Ergebnisse.

Das letzte Problem für die wissenschaftliche Betrachtung ist die Unzuverlässigkeit bezüglich der Kennzahlen. Teilweise wurden Kennzahlen ermittelt, die keinen Zusammenhang zu der Passwortsicherheit haben. Beispielweise hat die Informationsdichte keinen direkten Zusammenhang zur Schwierigkeit ein Passwort zu knacken.[9]

4. Passworterhebung und Anonymisierung

Die Passworterhebung erfolgt bei dem amerikanischen Unternehmen Yahoo!. Es bietet mit 12.200 Mitarbeitern zahlreiche Dienste, wie Email, Search uvm. für seine Kunden an. Weltweit hat Yahoo 800 Millionen Kunden.[10]

Die Passwörter werden mit Hilfe eines Proxy-Servers gesammelt. Dieser wird vor die Login-Server von Yahoo! platziert und hat auf diese Weise Zugriff auf die Benutzernamen und die dazugehörigen Passwörter in Klartext. Der Proxy-Server ist erforderlich, weil bei Yahoo! Passwörter nicht im Klartext gespeichert werden. Die Passwörter werden in der Datenbank als Hash gespeichert und haben zusätzlich einen individuellen Salt. Der Salt ist nicht geheim und bewirkt, dass gleiche Passwörter von unterschiedlichen Benutzern nicht den gleichen Hashwert in der Datenbank haben.

Der Proxy-Server hat die Aufgabe die Passwörter für die spätere Analyse zu speichern und die Daten mit zusätzlichen demographischen Daten anzureichern. Dabei soll beachtet werden, dass die Passwörter kryptographisch gespeichert und die zusätzlichen demographischen Daten anonym erhoben werden.

Im einfachsten Fall speichert der Proxy-Server das Passwort und die demographischen Zusatzinformationen, wie in der folgenden Abbildung dargestellt, in einer Liste. Der Benutzername ist nach Erhebung der Zusatzinformationen nicht weiter relevant. F1 und F2 stehen exemplarisch für Funktionen um bestimmte Zusatzinformationen zu ermitteln. [11]

[9] Vgl. Bonneau (2012), s. S. 3.
[10] Vgl. o. V. (2014), o. S.
[11] Vgl. Bonneau (2012), s. S. 6f.

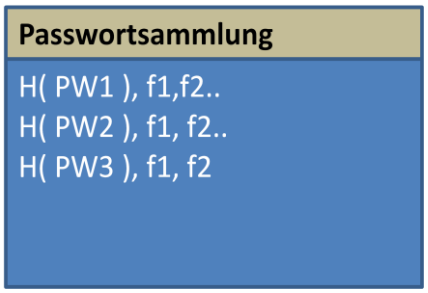

Abbildung 3: Exemplarische Passwortsammlung

Dabei gibt es folgendes zu beachten. Die Bildung des Hashwertes in der Form „H(PW1)“ ist nicht ausreichend, weil dann ein Angreifer einen Wörterbuchangriff gegen die Hashwerte durchführen könnte. Sollte der Benutzer durch seine demographischen Daten ebenfalls identifiziert werden können, dann hätte ein Angreifer den Benutzernamen und das Passwort. Aus diesem Grund wird eine schlüsselabhängige Hashfunktion verwendet[12]. Vor der Erhebung der Daten wird ein Schlüssel r erzeugt, der mit dem Passwort verknüpft wird. Anschließend wird ein Hashwert erzeugt H(r || Passwort). Der Schlüssel wird nach der Erhebung gelöscht, sodass ein Angriff auf den Hash-Wert erschwert wird.[13]

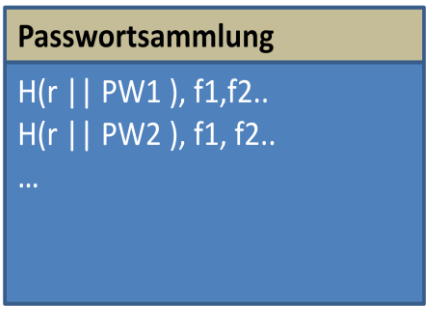

Abbildung 4: Exemplarische Passwortsammlung mit kryptographischer Hashfunktion

Die Speicherung aller Passwörter in einer Liste ermöglicht die Identifizierung von Benutzern, die ein gleiches Passwort verwenden. Es existieren also Gruppen von Benutzern, die ein Passwort verwenden. Dies ist problematisch, wenn für einen Benutzer der Gruppe das Passwort geknackt wird, weil dann für alle Mitglieder das Passwort bekannt ist. Das grundlegende Problem ist, dass mit Hilfe der demographischen Informationen die Möglichkeit besteht eine Re-Identifikation durchzuführen. Demographische Informationen sind beispielsweise das Geschlecht, das Alter oder auch die Verwendung von Diensten, wie Email und viele weitere. Diese Informationen stehen derzeit in einer Liste mit dem verschlüsselten Passwort. Um das Problem zu lösen wird pro Information eine Liste gepflegt, die lediglich aus einer Sammlung von Passwörtern besteht. Diese wird in der folgenden Abbildung dargestellt.

[12] Vgl. Schmeh (2013), s.S. 259ff.
[13] Vgl. Bonneau (2012), s. S. 7.

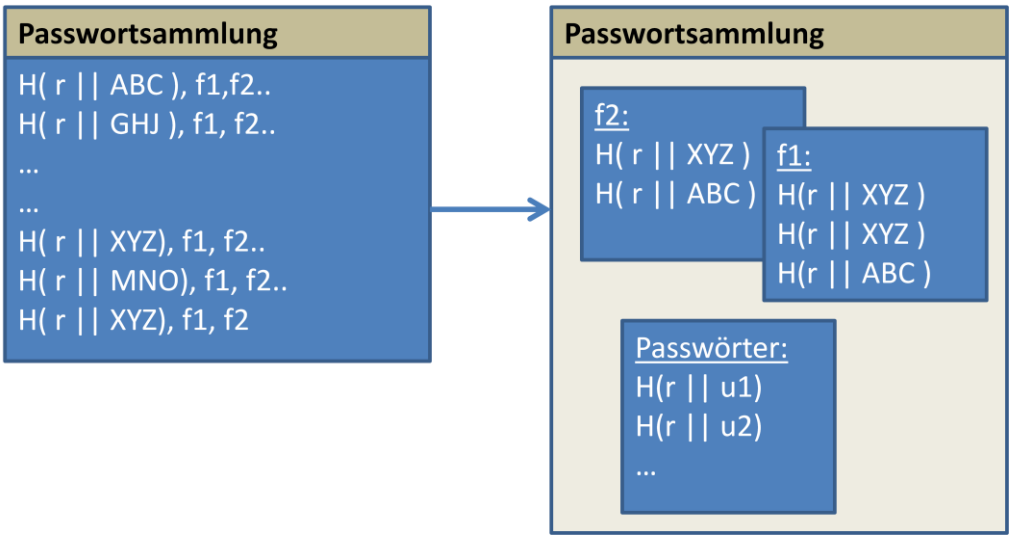

Abbildung 5: Exemplarische Passwortsammlung mit getrennten Listen

Auf diese Weise wird für Passwörter, die doppelt vorkommen, die Struktur eliminiert. Für einzigartige Passwörter funktioniert dies nicht, aber hier existiert auch nicht das Gruppenproblem. Damit keine Passwörter durch mehrfaches Login doppelt gewertet werden, wird eine weitere Liste mit den Benutzernamen angelegt, die im Anschluss der Erhebung nicht weiter benötigt wird.[14]

Die Datenerhebung wurde im Jahr 2011 im Zeitraum 23 Mai bis 25 Mai durchgeführt. In diesem Zeitraum haben sich 69.301.337 Accounts angemeldet und es wurden 328 Teilmengen gebildet.[15]

5. Kennzahlen

5.1. Zielsetzung

Im Vergleich zu den bisherigen Studien ermittelt diese Studie nicht die Erfolgsrate, sondern mit Hilfe von Kennzahlen eine sogenannte Guessing difficulty. Der Begriff Guessing difficulty soll Ausdrücken, wie leicht die Passwörter in einer Passwortsammlung erraten werden können. Die Kennzahlen basieren auf statistischen Formeln und werden nicht durch Konfigurationseinstellungen der Programme bzw. speziell optimierte Mangling-Rules verändert. Auf diese Weise werden Verzerrungen durch die Programme eliminiert. Man kann beispielsweise die Kennzahlen auf verschiedenen Passwortsammlungen anwenden und die Kennzahlen anschließend vergleichen.[16]

[14] Vgl. Bonneau (2012), s. S. 7.
[15] Vgl. Bonneau (2012), s. S. 7.
[16] Vgl. Bonneau (2012), s. S. 4.

Das Vorgehen für die Ermittlung der Guessing difficulty wird in der folgenden Abbildung verdeutlicht. Das Wörterbuch wird aus der Passwortsammlung generiert. Die Größe und welche Passwörter in dem Wörterbuch sind, hängt von der Kennzahl ab. Anschließend wird mit Hilfe des Wörterbuchs die Guessing difficulty berechnet. Das Wörterbuch entspricht einem Best-Case-Wörterbuch. Alle Passwörter sind dem Angreifer bekannt und so kann das Wörterbuch aus den häufigsten Passwörtern generiert werden. Dies ist ein wesentlicher Grund, weshalb die Kennzahlen vergleichbar und nicht von Wörterbüchern bzw. Mangling-Rules abhängig sind.[17]

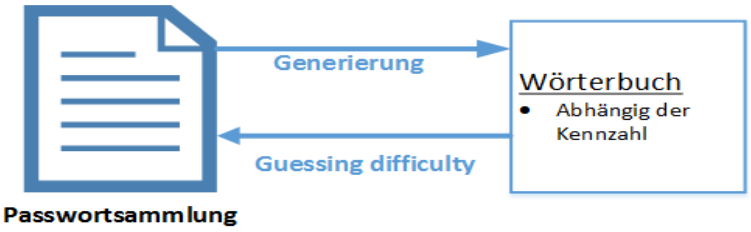

Abbildung 6: Anwendung der Kennzahlen

Die Generierung des Wörterbuchs aus der Passwortsammlung ermöglicht die Verwendung von anonymisierten Passwörtern. Wichtig ist, dass die Gleichheit von doppelten Passwörtern weiterhin erkennbar bleiben muss. Diese Bedingung wird durch die Passworterhebung sichergestellt. Desweiteren werden Passwortsammlungen mit eine großen Anzahl an Passwörtern genommen um mögliche Verzerrungen durch Besonderheiten in einer Passwortsammlung zu reduzieren.[18]

5.2. min-entropy

Die Kennzahl min-entropy gehört zur Renyi-Entropie, welche auf der Shannon Entropie basiert. Die Shannon Entropie beschreibt statistische Eigenschaften einer Nachrichtenquelle und wurde in bisherigen Studien häufig angewendet. Jedoch kann die Shannon Entropie für ein Passwort nicht die Komplexität der Information bzw. semantischen Aspekte einer Information ausdrücken.[19]

Die min-entropy ist eine Variante der Renyi Entropie und wird als Kennzahl für die Guessing difficulty verwendet. Die Kennzahl ist ein Worst-Case Szenario für das Erraten von Passwörtern. Der Angreifer versucht pro Account nur eine Passworteingabe und bricht dann das Erraten für den Account ab. Das gewählte Passwort für den einen Versuch ist das Passwort, das

[17] Vgl. Bonneau (2012), s. S. 4.
[18] Vgl. Bonneau (2012), s. S. 4.
[19] Vgl. Miller (2003), s. S. 88.

in der Passwortverteilung am Häufigsten vorkommt. Dieses Verfahren kann ebenfalls für Online-Attacken auf einen Account angewendet werden.[20]

Die Abbildung 7 zeigt das Vorgehen exemplarisch. In diesem Beispiel wird als Passwort PW10 genutzt. Ein typisches Beispiel für solch ein Passwort kann „123456“ sein. Im Ergebnis wird die Anzahl erratener Accounts betrachtet. Die Einheit ist „Bits“.

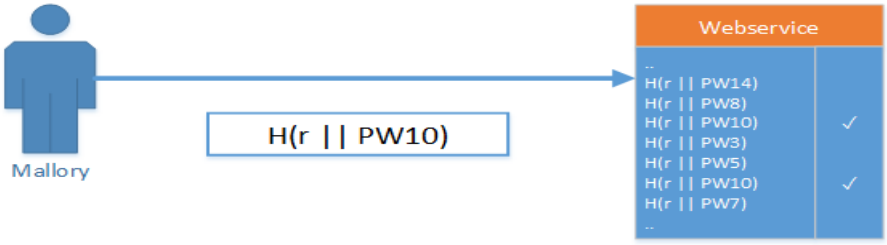

Abbildung 7: Exemplarische Anwendung min-entropy

5.3. β-succes-rate

Das Vorgehen für die Kennzahl β-success-rate entstammt dem Prinzip Passwörter exzessiv durchzuprobieren. Im englischen wird dies als guesswork bezeichnet. Bei dem Vorgehen werden pro Account alle Passwortvarianten durchprobiert. Das Verfahren scheitert bei Passwörtern, die nicht in einem Wörterbuch enthalten sind und für die keine Mangling-Rules existieren. Daher wurde dieses Verfahren weiterentwickelt und es erfolgt ein Abbruch nach einer bestimmten Anzahl an Versuchen pro Account. Diese Verfahren mit Abbruch werden als Partial guessing metrics bezeichnet. Das Ziel für den Angreifer ist es eine höhere Anzahl an Accounts zu knacken und dabei die Anzahl an Versuchen pro Account zu reduzieren.

Die Kennzahl β-success-rate bricht nach einer definierten Anzahl ab. Ursprünglich wurde sie von Boztas entwickelt. Die Abbildung 8 zeigt exemplarisch das Vorgehen. In diesem Fall ist die Anzahl auf 4 Passwörter festgelegt. Die 4 Passwörter sind die häufigsten Passwörter in der Passwortsammlung. Das Ergebnis ist die prozentuale Anzahl an gebrochenen Accounts.[21]

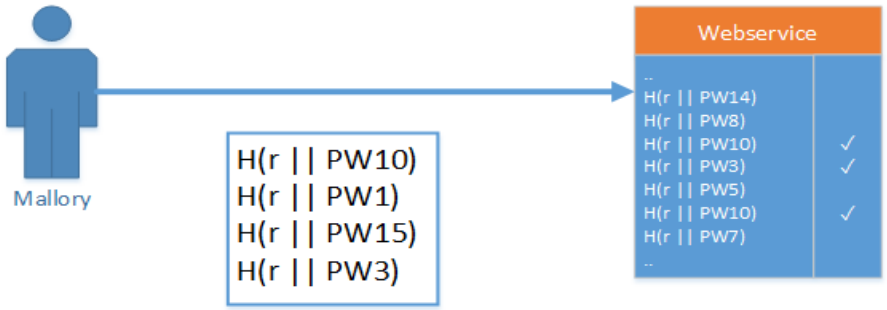

Abbildung 8: Exemplarische Anwendung β -success-rate

[20] Vgl. Bonneau (2012), s. S. 4.
[21] Vgl. Bonneau (2012), s. S. 5.

Das Verfahren eignet sich für Online-Attacken und die Anzahl an Versuche wird auf den Wert gesetzt, für den eine Webseite fehlerhafte Login-Versuche zulässt. In der Praxis hat sich jedoch gezeigt, dass einige Webseiten keine Beschränkungen bezüglich der Versuche haben.[22]

5.4. α-guesswork

Die Kennzahl α-guesswork basisert auf der Kennzahl α-work-factor, welche von Pliam entwickelt wurde. Die Kennzahl von Pliam ermittelt die Anzahl an Versuchen pro Account um einen gewünschten Anteil an geknackten Accounts zu erhalten. Sie wurde von Bonneau weiterentwickelt, weil in der ursprünglichen Form nicht berücksichtigt wurde, dass nach einem erfolgreichen Probieren für den Account vorzeitig abgebrochen werden kann. Die neue Kennzahl α-guesswork berücksichtigt dies. Diese Kennzahl eignet sich nicht für Online-Attacken, weil pro Account eine hohe Anzahl an Versuchen benötigt wird und daher wird die Kennzahl für Offline-Attacken verwendet.

Die Abbildung 9 zeigt die Anwendung der Kennzahl. In der Abbildung wird verdeutlicht, dass die Anzahl an Versuchen nicht festgelegt ist, sondern dass eine prozentuale Anzahl an geknackten Accounts erwartet wird. Der Angreifer benötigt ein Wörterbuch, welches aus einer Anzahl x Passwörtern besteht um beispielsweise 25% der Accounts zu knacken.[23]

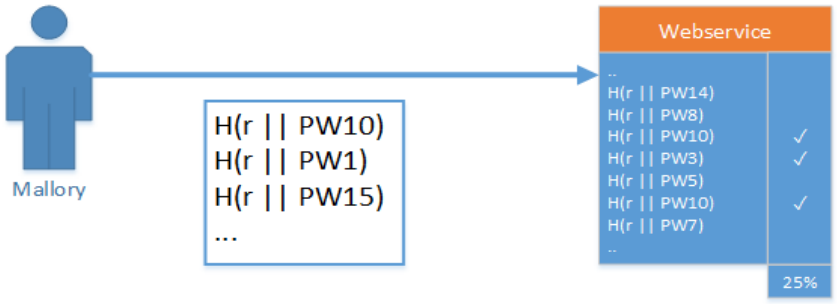

Abbildung 9: Exemplarische Anwendung α-guesswork

5.5. Konvertierung

Die Kennzahl β-success-rate und α-guesswork werden in die Einheit Bit konvertiert. Die Kennzahl min-entropy liegt bereits in dieser Einheit vor. Die Konvertierung der beiden Kennzahlen ermöglicht den Vergleich aller Kennzahlen und drückt die effektive Schlüssellänge aus. Die Angabe der effektiven Schlüssellänge ist in der Kryptographie üblich. Die Konvertierung erfolgt durch Anwendung des Logarithmus auf die Formeln β-success-rate und α-guesswork.[24] Die Abkürzung in den nachfolgenden Tabellen für die Kennzahl min-entropy ist

[22] Vgl. Bonneau (2012), s. S. 5.
[23] Vgl. Bonneau (2012), s. S. 5.
[24] Vgl. Bonneau (2012), s. S. 5.

H, für β-success-rate λ und für α-guesswork G. Wenn ein „^“ vor der Abkürzung verwendet wird, dann drückt dies aus, dass die Kennzahl in der Einheit Bit vorliegt.

6. Analyse Yahoo!-Passwortsammlung

6.1. Externer Vergleich

Die Passwortsammlung von Yahoo! wird zunächst mit anderen Passwortsammlungen verglichen. Es werden Passwortsammlungen verwendet, die den statistischen Ansprüchen genügen und verfügbar sind. Die Passwortsammlungen RockYou und Battlefiled Heroes sind durch sogenannte data leaks (Datenabflüsse) in die Öffentlichkeit gelangt. Der Leak von RockYou umfasst 32 Millionen Passwörter, während der Leak von Battlefield Heroes 548774 Passwörter enthält. Im Vergleich zeigt sich, dass ^H und $\hat{\lambda}_{10}$ um einen Bit abweichen. Bei $\hat{G}_{0.5}$ und $\hat{G}_{0.5}$ gibt es eine Abweichung von bis zu 2 Bits. Die Abweichung ist minimal und zeigt, dass die Kennzahlen in der Yahoo! Passwortsammlung vergleichbar sind, obwohl die Passwortsammlungen aus anderen Bevölkerungsteilen stammen.

	M	$\hat{H}_\infty$	$\hat{\lambda}_{10}$	$\hat{G}_{0.25}$	$\hat{G}_{0.5}$
Yahoo! (2011)	69301337	6.5	9.1	17.6	21.6
RockYou (2009)	32603388	6.8	8.9	15.9	19.8
Battlefield Heroes (2011)	548774	7.7	9.8	16.5	20.0

Abbildung 10: Vergleich unterschiedliche Passwortsammlungen[25]

Im Vergleich erreicht die Passwortsammlung von Yahoo! bezüglich der Kennzahlen min-entropy den kleinsten Wert. Die Kennzahl β-success-rate wird mit der Ausprägung 10 angewendet und Yahoo! erreicht einen mittleren Wert. Die Kennzahl α-guesswork liegt in den Ausprägungen 0.25 und 0.5 an geknackten Accounts vor. Hier erreicht Yahoo im Vergleich zu RockYou und Battlefield Heroes bessere Werte.[26]

6.2. Vergleich von Teilmengen

Bei der Datenerhebung wurden 328 Teilmengen erhoben. Als Teilmenge werden demographische Eigenschaften bezeichnet. Einige Beispiele für die Teilmengen werden in der Tabelle abgebildet.

In der Betrachtung der Kennzahlen aller Teilmengen innerhalb der Passwortsammlung von Yahoo! gibt es bezüglich min-entropy Abweichungen von 5-9.1 Bit. Bei β-success-rate gibt es

[25] Vgl. Bonneau (2012), s. S. 10.
[26] Vgl. Bonneau (2012), s. S. 10.

eine Abweichung von 7.5 bis 10.9 Bit und bezüglich der Kennzahl α-guesswork gibt es größere Abweichungen. Im Anhang ist die Tabelle vollständig hinterlegt. Im Folgenden werden Auszüge aus der Tabelle dargestellt.

Bei allgemeinen Informationen, wie beispielsweise dem Geschlecht gibt es die Ausprägungen männlich und weiblich. In den Kennzahlen wird deutlich, dass männliche Accounts schwächer gegen Online-Attacken (^H, ^λ), aber stärker gegen Offline-Attacken sind. Aber der Unterschied zwischen den beiden Ausprägungen liegt lediglich bei ca. 1 Bit. Auch beim Alter liegen die Abweichungen bezüglich der einzelnen Altersgruppen nur bei einem Bit. In der Tendenz zeigt sich, dass ältere Altersgruppen bessere Passwörter wählen. Bezüglich der unterschiedlichen Sprachen der Accounts werden die Abweichungen teilweise deutlicher. Besonders die Sprache Indonesisch hat bezüglich den Sprachen Deutsch und Koreanisch deutlich schlechtere Werte. Die Abweichung bezüglich min-entropy zwischen Indonesien und Korea liegt bei 4 Bit.[27]

	M	$\hat{H}_\infty$	$\hat{\lambda}_{10}$	$\hat{G}_{0.25}$	$\hat{G}_{0.5}$
all passwords	69301337	6.5	9.1	17.6	21.6
gender (self-reported)					
female	30545765	6.9	9.3	17.2	21.1
male	38624554	6.3	8.8	17.7	21.8
age (self-reported)					
13–24	18199547	6.3	8.7	16.7	20.9
25–34	22380694	6.2	8.8	17.1	21.2
35–44	12983954	6.8	9.4	17.4	21.3
45–54	8075887	7.3	9.8	17.3	21.3
>= 55	7110689	7.5	9.8	17.3	21.4
language					
Chinese	1564364	6.5	8.6	17.3	22.0
German	1127474	7.4	9.7	15.8	19.7
English	55805764	6.5	9.0	17.4	21.5
French	2084219	6.9	9.0	14.8	18.6
Indonesian	1061540	5.5	7.9	14.3	17.0
Italian	811133	6.8	9.0	14.5	18.0
Korean	530759	7.5	9.5	18.1	22.7
Portuguese	2060256	6.5	9.0	15.6	18.8
Spanish	3065901	6.6	9.1	15.6	19.7

Abbildung 11: Ausszug Vergleich Teilmengen 1[28]

Einige Teilmengen bieten die Möglichkeit Trends zu erkennen. Beispielsweise lässt sich über Benutzer mit häufigen Passwortänderungen feststellen, dass diese bessere Passwörter

[27] Vgl. Bonneau (2012), s. S. 10f.
[28] Vgl. Bonneau (2012), s. S. 12.

verwenden. Hier werden die Abstufungen kein Passwortwechsel, 1 Wechsel, mehr als einer und mindestens 5 Passwortwechsel unterschieden.

Die Tatsache, dass einige Benutzer schwache Passwörter benutzen, ist den Unternehmen bekannt und es gibt Bestrebungen die Benutzer bei der Generierung von besseren Passwörtern zu unterstützen. Dies ist auch bei Yahoo! erfolgt. In der Passwortsammlung von Yahoo! können Teilmengen unterschieden werden, welche bei der Anmeldung bei Yahoo! keine Passwortrestriktionen hatten und welche, bei denen die Passwortlänge mindestens 6 Zeichen lang sein musste und die Passwortstärke graphisch illustriert wurde. Die Maßnahmen zeigen kaum Änderungen bezüglich der Passwortsicherheit bei Online-Angriffen. Bei Offline-Angriffen gibt es eine schwache positive Tendenz.

Deutliche Tendenzen werden bei Benutzern deutlich, die sich von unterschiedlichen Standorten anmelden. Hier zeigt sich, dass die Passwortstärke mit der Anzahl unterschiedlicher Standorte zunimmt.[29]

	M	^H∞	^λ10	^G0.25	^G0.5
all passwords	69301337	6.5	9.1	17.6	21.6
	password requirements at registration				
none	20434875	6.6	9.2	16.8	20.7
6 char. minimum	13332334	6.5	9.0	17.6	21.6
	number of login locations				
1	16447906	6.0	8.6	17.1	21.1
>= 2	52853431	6.7	9.2	17.7	21.7
>= 10	17146723	7.3	9.7	18.3	22.6
	number of password changes				
none	52117133	6.2	8.8	17.1	20.9
1	9608164	8.3	10.4	18.8	23.2
> 1	7576040	8.6	10.7	19.5	24.2
>= 5	930035	9.1	10.9	19.7	25.9

Abbildung 12: Ausszug Vergleich Teilmengen 2[30]

Die Tabelle von Bonneau enthält einen Auszug für die 328 Teilmengen. Einige Ergebnisse sind nur textuell beschrieben. Beispielsweise ist erkannt worden, dass Benutzer mit einer Email-Recovery schwächere Passwörter wählen. Email-Recovery ermöglicht auf einfache Weise ein Passwort via Email zurückzusetzen.[31]

[29] Vgl. Bonneau (2012), s. S. 11.
[30] Vgl. Bonneau (2012), s. S. 12.
[31] Vgl. Bonneau (2012), s. S. 11.

6.3. Effekte durch Wörterbücher

Bisher wurden die Kennzahlen ermittelt, indem das Best-Case-Wörterbuch für eine Teilmenge angewendet wurde. In diesem Abschnitt werden gezielt Wörterbücher von fremden Teilmengen angewendet um die Effizienzverluste gegenüber dem Best-Caste-Wörterbuch zu messen. Dazu wird die Kennzahl β-success-rate mit der Anzahl 1000 genommen. Die Kennzahl wird nicht mit den anderen Kennzahlen verglichen und daher erfolgt keine Konvertierung in die Einheit Bit. Die Ergebnisse sind in der folgenden Abbildung dargestellt. Hier werden für die Teilmengen männlich und weiblich die Wörterbücher männlich und weiblich angewendet. Der Effizienzverlust gegenüber dem Best-Case-Wörterbuch liegt bei 10-15%.

		dictionary	
		♀	♂
target	♀	**7.8%**	6.8%
	♂	6.3%	**7.1%**

Abbildung 13: Vergleich Effizienz spezieller Wörterbücher bezüglich Geschlecht[32]

Der Effizienzverlust ist so gering, dass der Autor es für unwahrscheinlich hält, dass beispielsweise ein Angreifer sein Wörterbuch speziell auf männliche Benutzer abstimmen würde um einen männlichen Account zu knacken.

		dictionary												
		Chinese	German	Greek	English	French	Indonesian	Italian	Korean	Portuguese	Spanish	Vietnamese	global	minimax
target	Chinese	**4.4%**	1.9%	2.7%	2.4%	1.7%	2.0%	2.0%	2.9%	1.8%	1.7%	2.0%	2.9%	2.7%
	German	2.0%	**6.5%**	2.1%	3.3%	2.9%	2.2%	2.8%	1.6%	2.1%	2.6%	1.6%	3.5%	3.4%
	Greek	9.3%	7.7%	**13.4%**	8.4%	7.4%	8.1%	8.0%	8.0%	7.7%	7.8%	7.7%	8.6%	8.9%
	English	4.4%	4.6%	3.9%	**8.0%**	4.3%	4.5%	4.3%	3.4%	3.5%	4.2%	3.5%	7.9%	7.7%
	French	2.7%	4.0%	2.9%	4.2%	**10.0%**	2.9%	3.2%	2.2%	3.1%	3.4%	2.1%	5.0%	4.9%
	Indonesian	6.7%	6.3%	6.5%	8.7%	6.3%	**14.9%**	6.2%	5.8%	6.0%	6.2%	5.9%	9.3%	9.6%
	Italian	4.0%	6.0%	4.6%	6.3%	5.3%	4.6%	**14.6%**	3.3%	5.7%	6.8%	3.2%	7.2%	7.1%
	Korean	3.7%	2.0%	3.0%	2.6%	1.8%	2.3%	2.0%	**5.8%**	2.4%	1.9%	2.2%	2.8%	3.0%
	Portuguese	3.9%	3.9%	4.0%	4.3%	3.8%	3.9%	4.4%	3.5%	**11.1%**	5.8%	2.9%	5.1%	5.3%
	Spanish	3.6%	5.0%	4.0%	5.6%	4.6%	4.1%	6.1%	3.1%	6.3%	**12.1%**	2.9%	6.9%	7.0%
	Vietnamese	7.0%	5.7%	6.2%	7.7%	5.8%	6.3%	5.7%	6.0%	5.8%	5.5%	**14.3%**	7.8%	8.3%

Abbildung 14: Vergleich Effizienz spezieller Wörterbücher bezüglich Sprache[33]

Die Abbildung 14 zeigt für verschiedene Sprachen den Effizienzverlust durch Anwendung von fremden Wörterbüchern. Es zeigt sich, dass der größte Effizienzverlust entsteht, wenn man französische Passwörter mit einem vietnamesischen Wörterbuch angreift. In der Tabelle sind zwei spezielle Wörterbücher hinzugefügt. Das globale Wörterbuch enthält die häufigsten Passwörter der gesamten Passwortsammlung. Das Wörterbuch „minimax“ ist speziell angepasst,

[32] Vgl. Bonneau (2012), s. S. 12.
[33] Vgl. Bonneau (2012), s. S. 12.

sodass es gegen alle Teilmengen effektiv ist.[34] Das globale und das “minimax” Wörterbuch funktionieren grundsätzlich gegen alle Teilmengen der Tabelle gut.[35]

7. Fazit

Wie gezeigt bietet ein Passwort eine Sicherheit von ca. 10 Bit gegen Online-Angriffe, bei denen ein Angreifer 10 Angriffe pro Account durchführen kann. Bei einem Offline-Angriff, bei dem ein Angreifer die Hälfte der Accounts brechen möchte, bietet ein Passwort eine Sicherheit von 20 Bit. Die Voraussetzung dafür ist, dass der Angreifer ein optimales Wörterbuch verwendet.[36]

Auch wenn die Passwortrestriktionen kaum Erfolg in dieser Studie gezeigt haben, bieten diese für Sicherheitsforscher gute Ansatzmöglichkeiten. Nur die Mindestlänge festzulegen reicht nicht aus, sondern es müssen Komplexitätsregeln für die Passwörter gefordert werden. Auch die Verhinderung von einfachen Passwörtern kann die Werte für die Kennzahlen deutlich erhöhen. Ein großes Problem bleibt, dass ein Benutzer nicht einschätzen kann, wie leicht ein Passwort zu erraten ist.

[34] Vgl. Bonneau (2012), s. S. 13.
[35] Vgl. Bonneau (2012), s. S. 13.
[36] Vgl. Bonneau (2012), s. S. 13.

8. Literaturverzeichnis

- Bonneau, J (2012)

The science of guessing: analyzing an anonymized corpus of 70 million passwords,

http://www.jbonneau.com/doc/B12-IEEESP-analyzing_70M_anonymized_passwords.pdf,

Stand: 5.1.2014,

San Francisco, 2012.

- Miller, M (2003)

Symmetrische Verschlüsselungsverfahren, 1. Auflage, Wiesbaden, 2003.

- o.V. (2014)

Yahoo! 2013 Annual Report,

https://investor.yahoo.net/annuals.cfm,

Stand: 5.1.2014.

- Schmeh, K (2013)

Kryptographie, 5. Auflage, Heidelberg 2013.

III Anhangsverzeichnis

IV Anhang

IV1 Vergleich Teilmengen

	M	$\hat{H}_\infty$	$\hat{\lambda}_{10}$	$\hat{G}_{0.25}$	$\hat{G}_{0.5}$
all passwords	**69301337**	**6.5**	**9.1**	**17.6**	**21.6**
gender (self-reported)					
female	30545765	6.9	9.3	*17.2*	21.1
male	38624554	6.3	8.8	17.7	*21.8*
age (self-reported)					
13–24	18199547	6.3	8.7	16.7	20.9
25–34	22380694	6.2	8.8	*17.1*	21.2
35–44	12983954	6.8	9.4	*17.4*	21.3
45–54	8075887	7.3	9.8	*17.3*	21.3
$\geq$ 55	7110689	7.5	9.8	*17.3*	21.4
language preference					
Chinese	1564364	*6.5*	8.6	*17.3*	*22.0*
German	1127474	7.4	9.7	15.8	19.7
English	55805764	6.5	*9.0*	*17.4*	21.5
French	2084219	6.9	*9.0*	14.8	18.6
Indonesian	1061540	5.5	7.9	14.3	17.0
Italian	811133	6.8	*9.0*	14.5	18.0
Korean	530759	7.5	9.5	18.1	22.7
Portuguese	2060256	6.5	*9.0*	15.6	18.8
Spanish	3065901	6.6	*9.1*	15.6	19.7
tenure of account					
$\leq$ 1 y	5182527	6.9	*9.1*	18.0	22.5
1–2 years	5182527	6.9	*9.1*	18.0	22.5
2–3 years	12261556	6.2	8.6	17.7	*21.8*
3–4 years	10332348	6.2	8.8	*17.5*	21.6
4–5 years	9290840	6.1	8.8	*17.2*	21.2
$\geq$ 5 years	29104856	6.8	9.3	*17.2*	21.2
password requirements at registration					
none	20434875	6.6	9.2	16.8	20.7
6 char. minimum	13332334	6.5	*9.0*	*17.6*	*21.6*
last recorded login					
$<$ 30 days	32627777	*6.5*	*9.0*	*17.5*	21.5
$<$ 90 days	55777259	6.5	*9.0*	*17.5*	21.5
$>$ 90 days	8212643	7.0	9.5	17.7	*21.9*
number of login locations					
1	16447906	6.0	8.6	*17.1*	21.1
$\geq$2	52853431	6.7	*9.2*	17.7	*21.7*
$\geq$ 10	17146723	7.3	9.7	18.3	22.6
number of password changes					
none	52117133	6.2	8.8	*17.1*	20.9
1	9608164	8.3	10.4	18.8	23.2
$>$1	7576040	8.6	10.7	19.5	24.2
$\geq$5	930035	9.1	10.9	19.7	25.9
number of password resets (forgotten password)					
none	61805038	6.4	8.9	*17.3*	21.3
1	4378667	8.2	10.5	19.2	23.8
$>$1	3117632	8.7	10.8	19.7	24.6
$\geq$5	387469	8.7	10.6	19.9	26.6
amount of data stored with Yahoo!					
1st quartile	9830792	5.6	8.2	*17.3*	21.5
2nd quartile	20702119	6.3	8.8	*17.5*	21.5
3rd quartile	21307618	6.8	9.3	*17.5*	21.4
4th quartile	17447029	7.6	10.0	17.8	*22.0*
usage of different Yahoo! features					
media sharing	5976663	7.7	10.1	18.0	*22.3*
retail	2139160	8.8	10.5	16.8	21.4
webmail	15965774	6.3	8.8	*17.4*	21.2
chat	37337890	6.2	8.7	*17.1*	21.2
social networking	14204900	7.1	9.6	17.7	*21.8*
mobile access	20676566	6.7	9.3	*17.1*	21.1
Android client	1359713	8.3	10.3	*17.3*	21.5
iPhone client	6222547	8.1	10.1	*17.6*	*21.6*
RIM client	3843404	7.6	10.0	*17.2*	21.1

[37]

[37] Bonneau (2012), s. S. 12.