Lösenordens säkerhetsrisker inom digitala system

Anton Winquist 2001166
Lösenordens säkerhetsrisker inom digitala system
Handledare: Jan Westerholm
Fakulteten för naturvetenskaper och teknik
Åbo Akademi
2023

Innehåll

Abstrakt

Personlig autentisering inom digitala system är någonting som många använder varje dag i någon form, vare det i form av lösenord, pinkoder eller någonting annat. Antalet lösenordsskyddade enheter eller konton per person är svår att estimera, och användaren själv vet ofta inte hur stort hens digitala fotavtryck är. Det enda säkra är att det globala antalet lösenords skyddade enheter och konton växer ständigt. System säkrade med lösenord är därför priviligierade mål för angripare, och antalet dataintrång under de senaste åren bevisar detta. Ett lösenord är ofta det enda som skyddar användarens kritiska system eller konto. Grundproblemet med lösenord är att de är användar skapade. Den vanliga användaren har inte tid att skapa, eller komma ihåg starka och unika lösenord för varje sida eller enhet som hen tar i användning. Är användarnas säkerhetsvanor gällande autentisering tillräckligt starka för att hålla deras konton och enheter säkra? Denna avhandlings uppgift är att få en person som har datateknologi som hobby, utbildning eller yrke medveten om säkerhetsriskerna förknippade med de lösenord hen använder och hur man kan förbättra situationen. Syftet är att bättre förstå lösenordets starkheter och svagheter, samt att upplysa användaren om hur lösenord står sig i jämförelse med andra autentiseringsmetoder .

Avhandlingen kommer att ge en kort bakgrund på lösenordet, analysera studier och användarstatistik av lösenord. Sammanfatta den största lösenordsanalysen som är baserad på ett dataset på 3.9 miljarder kränkta användarkonton från databasen från Have i been pwned. Statistiken används för att identifiera traditionella lösenordsmönster och för att få användaren att undvika dessa.

Kapitel 1

Inledning

Det finns flera olika sätt att autentisera sig till ett digitalt system. Allmänt sett så kan digital personlig autentisering klassas in i tre olika fundamentala kategorier, nämligen nn-ågonting man vet", någonting man haröch någonting man är". Någonting man vet är t.ex lösenord, pin-koder, mönster etc. Någonting man har är något fysiskt. Det kan vara t.ex en nyckel, ett kort, eller en smarttelefon med tvåfaktorsautentisering. Någonting man är representerar biometriska egenskaper som fingeravtryck, ansikten, ögon som gör en unik. En kombination av dessa gör processen märkbart starkare. [12]

I denna avhandling kommer fokuset att ligga på lösenord som hör till Någonting man vetdelen. Lösenord är fortfarande en av de mest använda autentiseringsmetoderna idag och används av miljontals användare dagligen.

//TODO: add more here

Kapitel 2

//TODO: hur fixar ja strukturen på detta??

2.1 Metod

En ?!systematisk?! genomgång av artiklar och studier associerade med lösenord och lösenordens säkerhetsrisker.

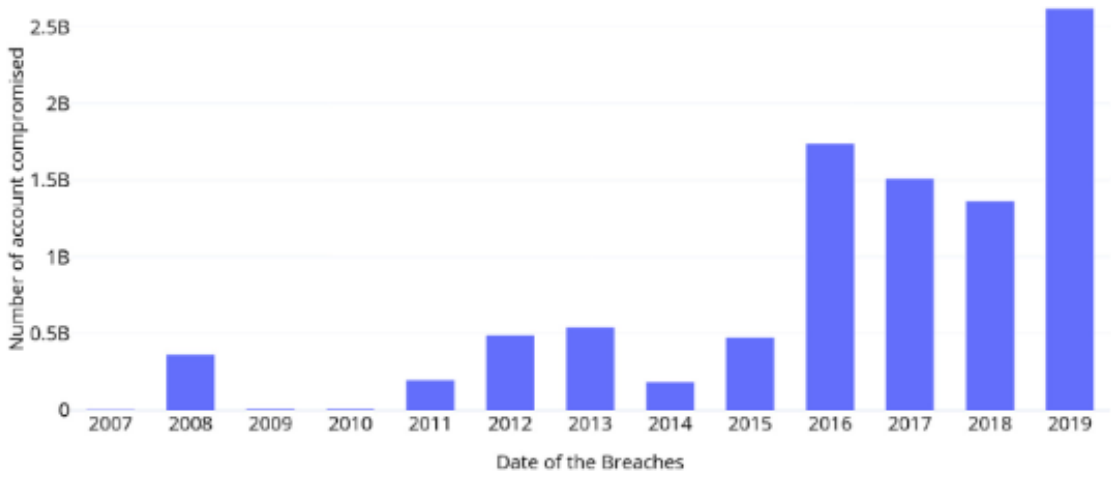
2.2 Historia

Textbaserade lösenord identifierades som en svag punkt för informationssystemens säkerhet redan år 1979. Robert Morris och Ken Thompson från Bell Laboratories utförde ett experiment för att försöka förstå användarnas lösenord och det visade sig att majoriteten av användarnas lösenord (86% av dem) var för svaga. Experimentets dataset bestod av 3 289 lösenord samlade från flera användare över en lång tidsperiod. Lösenorden var ofta för korta och bestod av enbart siffror eller små bokstäver. Utöver detta kunde många lösenord enkelt hittas i vanliga ordböcker. [11][9]

Trots den tidiga upplysningen om säkerhetsriskerna förknippade med lösenord har de varit och är fortfarande den vanligaste autentiseringsmetoden för datorsystem. Trots lösenordens betydelse som den första försvarslinjen i de flesta informationsystem, har lite uppmärksamhet ägnats åt deras säkerhetsegenskaper i praktisk användning. [8][11]

2.3 Dataläckor och data

Eftersom lösenordsäkrade system och konton omfattar en så stor andel är de privilegierade mål för angripare, och antalet dataintrång på de senaste åren bestyrkar detta. Mängden läckt användardata ökar till följd av den ökande frekvensen av dataintrång. Webbtjänsten Have i been pwned är skapad av Troy Hunt och den samlar ihop data läckt i dataläckor för att kunna informera både folk och företag om deras lösenord eller konton har blivit läckta. Syftet med hemsidan är att minska återanvändningen av lösenord och att hindra stulna inloggningsattacker (Eng: *Credential stuffing attacks).* Figur [INSERT] illustrerar det årliga antalet läckta konton som tilläggs till Have i been pwneds databas. Dataset version 5 innehåller 515 680 539 lösenord som korrelerar till 3 951 907 330 äkta användarkonton. Lösenordsåteranvändning är den drivande faktorn till den stora variationen i antalen. Med hjälp av dessa 3.9 miljarder kontons lösenordsuppgifter har forskare identifierat lösenordens vanligaste mönster. Analysen av dessa lösenord hjälper användare att lära sig om vilka mönster som gör ett lösenord svagt och därmed hur man kan skapa starkare lösenord.[7][8]



Figur 2.1: Antal läckta konton per år tillaggda i Have i been pwned

2.4 Mängden konton

På grund av den snabba tillväxten och ökande popularitet av internet teknologi, e-handel och det ökande antalet onlinetjänster som kräver lösenordsbaserad autentisering ökar måste användarna underhålla allt flera konton. Traditionellt rekommenderades det att man skulle minnas alla sina lösenord men då mängden lösenord man behövde ökade blev detta allt svårare. [6] Enligt en studie gjord av NordPass visade det sig att användare har i medeltal över 100 lösenord. För en vanlig användare är det opraktiskt att vara tvungen att

komma ihåg så många lösenord. [14]. Även lösenordets skapare Fernando Cobato meddelade i en intervju med The Wall Street Journal år 2014 att lösenorden har blivit något av en mardröm. Han meddelar att det är svårt att hålla reda på alla lösenord man har skapat och att man då lämnas med två val. Antingen så skriver man dem på ett papper, eller så använder man ett program som hanterar lösenorden. [6]

2.5 Hur användarens bakgrund påverkar på lösenordsval

Undersökningar har visat att en persons bakgrund påverkar valet av lösenord. En betydande faktor som påverkar lösenordsskapandet är demografiska faktorer som ålder, kön och nationalitet.

Resultatet från en 2004 studie riktat till studerande på Brown University visade att majoriteten av eleverna använde personlig information för att skapa lösenord. Studerandena tenderade att använda information som t.ex deras egen födelsedag, deras eget namn eller namn på deras vänner eller släktingar.[8, s. 2] Det har också visats att användarnamnet eller e-posten användaren väljer till sitt konto kan ge information om vad personen använder för lösenord.

Användares övertygelser och ofta missuppfattningar om vad som gör ett lösenord säkert kan också förklara deras lösenordsval. Användare tror ofta att lösenord som är svåra att stava är starkare, eller att säkerheten ökar då man tillägger specialtecken som "!"i slutet på lösenordet. [8]

Demografiska faktorer som ålder, kön och nationalitet påverkar lösenordsskapande. En 2012 analys på ca. 70 000 000 lösenord undersökte ifall gissbarheten för lösenorden ändrades när specifika ordböcker inriktade på specifika demografiska folkgrupper valdes ut. Det observerades att framgångsfrekvensen för en gissningsattack med 1000 gissningar, när en specificerad-till-en-specifik-befolkningsgrupp ordbok valdes presterade något bättre i kategorier som hade att göra med ålder, språk och tjänsteanvändning än en generisk ordbok. År 2015 visade det sig att desto närmare ordboken är semantiskt till målpersonen, desto högre var sannolikheten att lösenordet knäcktes. [8] Demografiska faktorerna i lösenord undersöktes år 2013 på ett universitet. Resultatet från studien visade att då man jämför lösenord mellan könen skapade män lite starkare lösenord än kvinnor. Då analyserade resultaten på basen av studieinriktning visade det sig att datavetenskaps studenter hade de starkaste lösenorden medan ekonomistuderanden hade bland dom svagaste.

Wang et al. (2017) [13] kollade på 12 olika dataset från specifika intressegrupper (Eng: communities) och undersökte om typen av webbsida lösenordet kom ifrån spelade någon

roll i lösenordsvalen. Tjänsten som lösenordet används på visade sig påverka lösenordsvalet. Till exempel var *jesus1* bland det vanligaste lösenordet till en Kristendomsfokuserad webbsida. I en dataläcka från *mangatraders.com* läcktes 881 468 konton (618 237 unika lösenord). Det visade sig att de top 100 vanligaste lösenorden representerade 4.76% av alla konton. Av dessa lösenord var 37.6% mangarelaterade (eller 1.79% av totala mängden). Detta stöder antagandet om att användare är inspirerade av webbsidans syfte och innehåll då de skapar sina lösenord.[8]

Flera lösenordsanalyser fokuserar på personer som har engelska som modersmål vilket inte visare hela bilden eftersom stor del av lösenordsanvändarna har andra modersmål. Resultatet från en analys på lösenordsfällor år 2014 Wang et al (2018) visade att runt 36.95% till 51.43% av personer med Kinesiskt modersmål använde personlig information för att skapa lösenrod, medan den siffran varierar från 12.76% till 29.94% för personer med Engelska som modersmål. Personer med Kinesiska som modersmål tenderade att använda numror och speciellt datum i deras lösenord oftare än personer med Engelska som modersmål. Användningen av siffror i lösenord är populärt i Asiatiska länder, troligen på grund av att de kan matas in digitalt lättare än ideogram, särskilt på Mobilenheter.

Det har också visat sig att toppmoderna lösenordsstyrkeestimatorer ofta överskjuter den upplevda komplexiteten när lösenord är sammansatta av icke-engelska ord och fragment. Estimationerna kan förbättras med att träna modellen med contextuellt relevant information. [3]

2.6 lösenordspolicyn

Som tidigare nämnt var Morris och Thompson de första som tog upp problemet med lösenordssäkerhet år 1979 och hur majoriteten av användare har för svaga lösenord. Ungefär 10 år senare ansåg man att för att maximera svårigheten att knäcka lösenord och förhindra snabba och simpla attacker skulle Lösenordspolicyn implementeras för att lösenorden skulle uppnå en viss mängd entropi. [11]

lösenordspolicyn är regler som hindrar användare att välja för lätta lösenord och ser därmed till att lösenord uppnår en viss nivå av entropi. Vanliga lösenordspolicyn som t.ex 2019 rekomendationerna från NIST (National Institute of Standards and Technology) innehåller bland annat krav på minimilängd på 8-64 tecken, inga ledtrådar, minimi antal försök på 10, förbud att använda upprepande eller sekventiella tecken, förbud av användningen av kontext-specifika ord (som namnet på tjänsten eller användaren) och tillåtelsen att använda alla ASCII/UNICODE tecken såväl som emojis och mellanrum. [4]

En tillräcklig lösenordskomplexitet (användningen av både stora och små bokstäver såväl som specialtecken) har länge rekomenderats som policy men det har visats att de orsakade mindre säkra människobeteenden istället för att förstärka säkerheten. Användningen av mera komplexa lösenord gör dem svårare för attackerare att knäcka med samtidigt gör dem svårare för användare att komma ihåg. För att undvika frustration tenderar användare att skriva ner deras användaruppgifter på t.ex en lapp vid deras bord eller beslutar att periodiskt använda samma eller liknande lösenord. En annan policy som gick igenom ett likadant öde var policyn för lösenordsutgång som numera också rekommenderas emot.

[4] Organisationer kan även använda sig av existerande blocklistor för att hindra användare att både skapa och använda osäkra lösenord. En implementering av detta är "Have i been pwned":s API *(Eng: Application Programming Interface)* som möjliggör att organisationer kan granska ifall deras användares E-post, användarnamn eller lösenord har blivit läckta i dataläckor. Have i been pwned's databas uppdateras regelbunder då nya dataläckor sker för att konstant vara uppdaterad. [7][8]

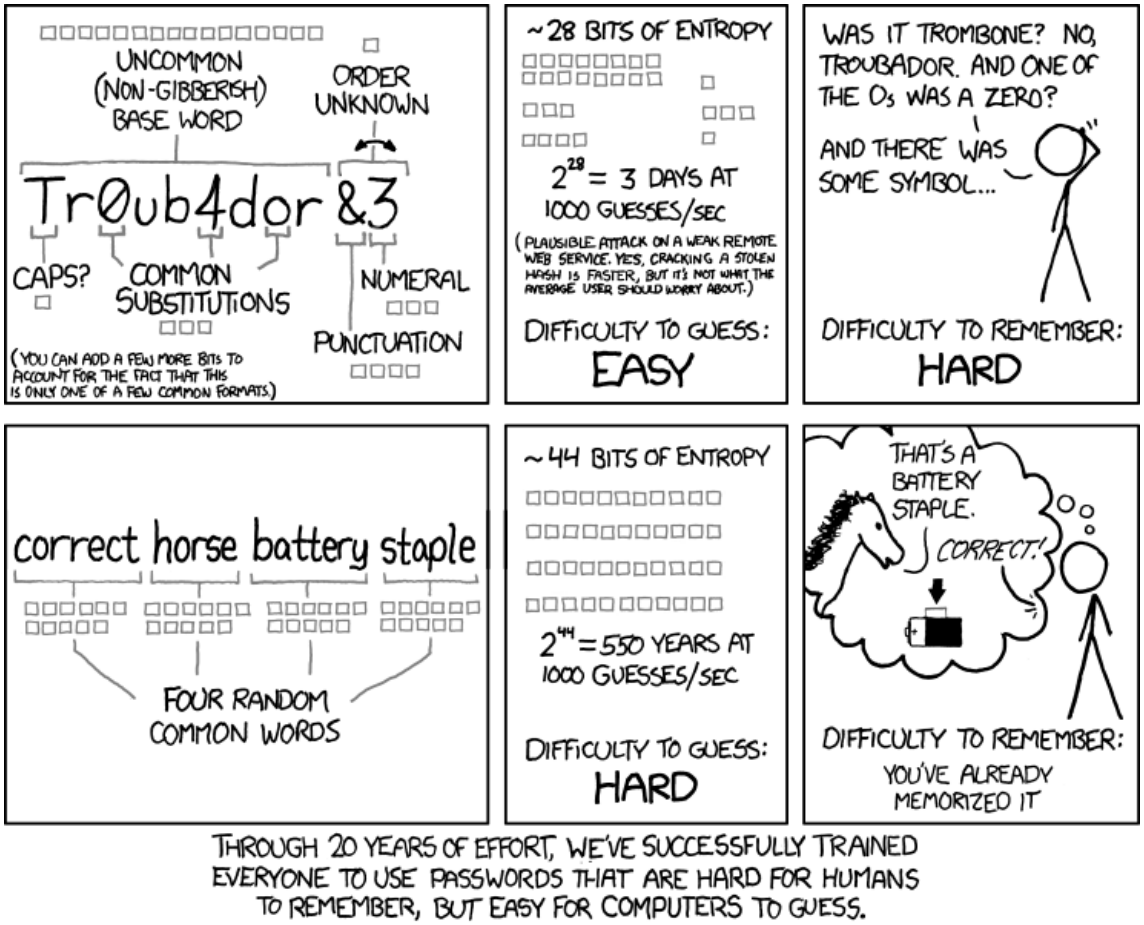
Striktare Lösenordspolicyn får användarna att spendera mera tid att välja lösenord samt att använda flera tecken. Lösenords som följer strikta lösenordspolicyn förblir ofta sårbara eftersom lösenordspolicyns krav uppnås på ett förutseebart sätt. Fast lösenordspolicyn uppmuntrar användarna att skapa och använda starkare lösenord visar Literatur att användare inte följer rekommendationer eller policyn. Det händer ofta eftersom policyn ofta är för strikta för att matcha användarnas förmågor. [11]För strikta policyn belastar ofta användarna alltför mycket, vilket leder till frustration. [10] Utan grundlig förklaring och användarutbildning har lösenordspolicyn nästan ingen effekt. [11]

Det har visat sig att lösenordslängden och användbarheten var inte alltid inverst proportionella.

Lösenordspolicyn är inte någonting som den enskiljda användaren kan påverka i första hand, men det är ändå vikgit att vara medveten om hur policyn påverkar lösenord.

2.7 Lösenfraser

Lösenord som består av en sekvens ord eller annan text kallas för en lösenfras (Eng: passphrase). Lösenfraser är som längre lösenord, med en längd som ofta är över 14 tecken. Eftersom lösenfraser ofta innehåller flera tecken gör det att entropin ökar, och därmed att tiden det tar för en dator att gissa i en så kallad brute-forceättack ökar märkbart. [10][11] Då användare själv fick välja lösenfraser följde distributionen det naturliga språket, vilket



Figur 2.2: xkcd tecknad serie. [15]

ökade gissbarheten jämfört med slumpmässigt skapade lösenfraser. [8]

Lösenfraser argumenteras vara starkare och lättare att komma ihåg än Lösenord. Både lösenord och lösenfraser har möjligheten att vara mycket starka, men ett längre lösenord anses vara svårare att komma ihåg än en längre lösenfras.[5]

För tre decennier har akademisk litteratur ansett att lösenfraser har potential att vara starkare och lättare att komma ihåg än korta lösenord, även då användbarheten inte blivit väl undersökt mot lösenord. En debatt har återupplivats av XKCDs webbserie (figur 2.1) [15] som föreslår att lösenfraser är ett alternativ för komplexa lösenordspolicyn. Webbserien har fått uppmärksamhet och t.ex använts för att hjälpa användare att skapa starka och minnesvärda lösenord. Webbserien har använts som inspiration och påverkat flera Amerikanska universitet att implementera lösenfraser i deras lösenordspolicyn. Trots brist på impiriska bevis på att lösenfraser skulle vara överlägsen rekommenderas de av vissa systemadministratörer och olika webbsidor. Undersökningar har upptäckt att användarvalda lösenfraser har visat sig vara lättare att gissa än vad man tidigare förväntat sig, vilket

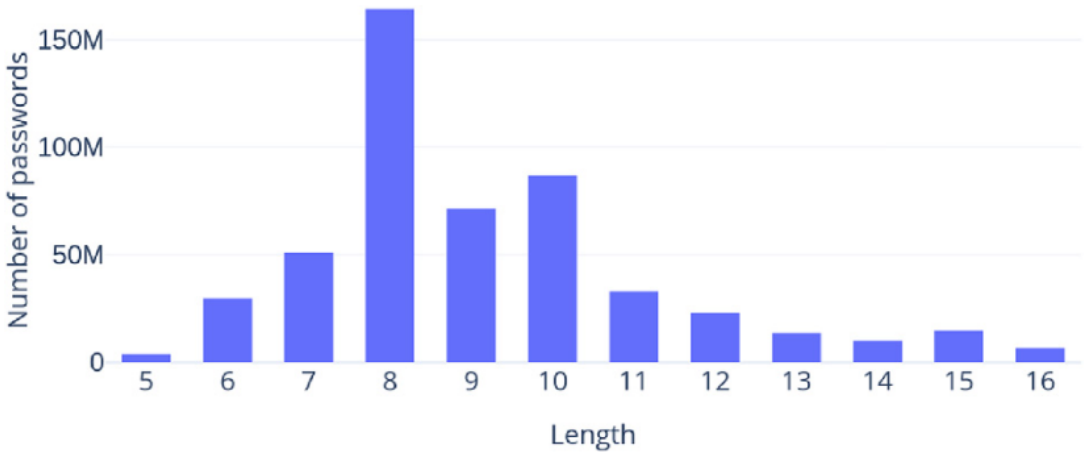
stöder undersökningen av datorgenererade lösenfraser som alternativ. [10]

Lösenfraser som bestod av 3-4 datorgenererade ord visade sig ha liknande grader av glömska som vanliga lösenord, men hade i allmänhet flera inmatningsfel.[8][11] Resultaten från en 2012 webbstudie med 1476 deltagare visade att lösenfraser och lösenord glöms i liknande skalor, orsakar liknande nivåer av svårighet och användarfrustration, och blev nedskrivna av majoriteten av deltagarna. [10]

Kapitel 3

Lösenordsanalys

3.1 lösenordens längd



Figur 3.1: De vanligaste lösenordslängderna

Figur (//TODO ADD HERE) ger en överblick på längdsdistributionen från de 515 680 539 lösenord från HIBP_v5. Den ögonfallande statistiken är att över 30% av de unika lösenorden är 8 tecken långa. En mycket sannolik förklaring till detta är att de flesta rekommendationer och policyn specifierar ett minimilängds krav, såsom minimum kravet på 8 tecken som NIST rekomenderar.[4] Den nästvanligaste längden är 10 som utgör 17% av lösenorden. 84% av lösenorden faller inuti intervallet 6-12 tecken. [8]

3.2 vanligaste lösenord

Figur (//TODO ADD HERE) visar de 20 vanligaste användargenererade lösenorden från HIBP_v5 med procent som indikerar hur stor andel konton som lösenordet är associerat med. Flera av lösenorden kan hittas bland de vanligaste eller sämsta lösenorden. Som man ser från tabellen är flera lösenord sekvenser av numror eller bokstäver som finns nära varandra på tangentborrdet. Denna typ av mönster kallas för tangentbordspromenad eller tangentbordsvandring (Eng: Keyboardwalk) och är de populäraste mönstret hittat i datasettet. [8]

Top 20 passwords in HIBP_v5.

Password	% of Total Accounts
123456	0.596%
123456789	0.197%
Qwerty	0.099%
password	0.094%
111111	0.079%
12345678	0.074%
abc123	0.072%
1234567	0.064%
password1	0.061%
12345	0.060%
1234567890	0.057%
123123	0.056%
000000	0.050%
iloveyou	0.041%
1234	0.033%
1q2w3e4r5t	0.030%
qwertyuiop	0.028%
123	0.026%
monkey	0.025%
Dragon	0.025%

Figur 3.2: De 20 vanligaste lösenorden

3.3 Lösenordskomplexitet

Tecken kan generellt delas in i 4 olika kategorier: Versaler, Gemener, siffror och specialtecken. PACK (Password Analysis and Cracking Kit)[1] analyserar lösenordens unika komposition och klassifierar dem enligt vilka karaktärset de använder. T.ex skulle ett lösenord som tillhör kategorin *loweralphaspecialnum* innehålla enbart små bokstäver, specialtecken och nummror oavsett i vilken ordning eller frekvens de förekommer. Ett exempel på ett lösenord som skulle kunna tillhöra denna kategori är: pa$$w0rd.

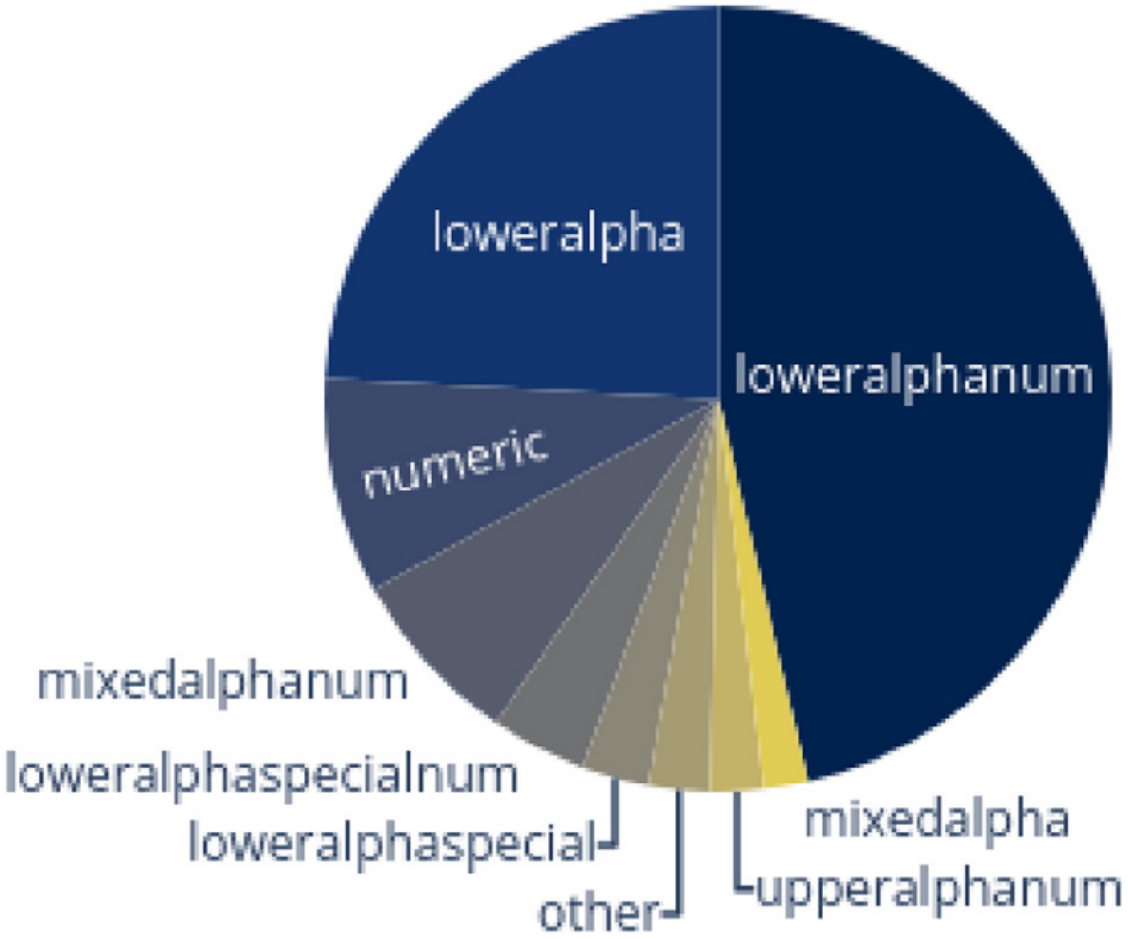
Figur [insert] illustrerar uppdelningen av lösenorden från datasettet HIBP_v5 med hjälp av kategoriseringen från PACK. Som man ser från cirkeldiagrammet hör 46% av lösenorden till kategorin *loweralphanum* vilket betyder att de enbart består av små bokstäver och siffror. De näststörsta och tredjestörsta kategorierna består av enbart små bokstäver (24%) och siffror (8%) respektive.

En märkbar observation från denna analys är att över 75% av alla 515 680 539 lösenord innehåller varken specialtecken eller stora bokstäver. Detta är inte oförväntat eftersom de flesta lösenordspolicyn kräver användningen av minst två olika karaktärset i lösenordet.

Analysen ovan kan finslipas eftersom den fokuserar enbart på enskilda tecken-set utan att ta i hänsyn den interna lösenordsstrukturen. Exempelvis lösenord som hör till kategorin loweredalphanum kunde innehålla lösenord som 12password, password12 och pass12word. En mera finslipad klassifikation där den interna tas i hänsynt skulle separera dessa till 3 olika kategorier. Lösenorden ovan kunde klassifieras till *digitstring, stringdigit* och *stringdigitstring*. Vidare klassifikationerna av lösernord är viktiga eftesrom metoden för att knäcka dom är olika. De olika klassifikationerna för lösernod kallas för maskar (Eng: *masks*). De 15 vanligaste maskar från dataset HIBP_v5 visualiseras i figur [INSERT].

Den vanligaste masken är *stringdigit*. Lösenord som tillhör denna kategori består av en sträng (stora och/eller små bokstäver) direkt följt av en siffra. Användare väljer ofta ett lösenord som består av en sträng (vanligtvis ett ord eller ett namn) och lägger sedan till siffror för att uppnå lösenordspolicyns krav på användningen av två eller flera karaktär-set och en tillräcklig längd. De tre nästvanligaste kategorier är strängar, siffror och siffror följt av en sträng. Dessa fyra kategorier representerar över 75% av lösenorden från datasettet. [8]

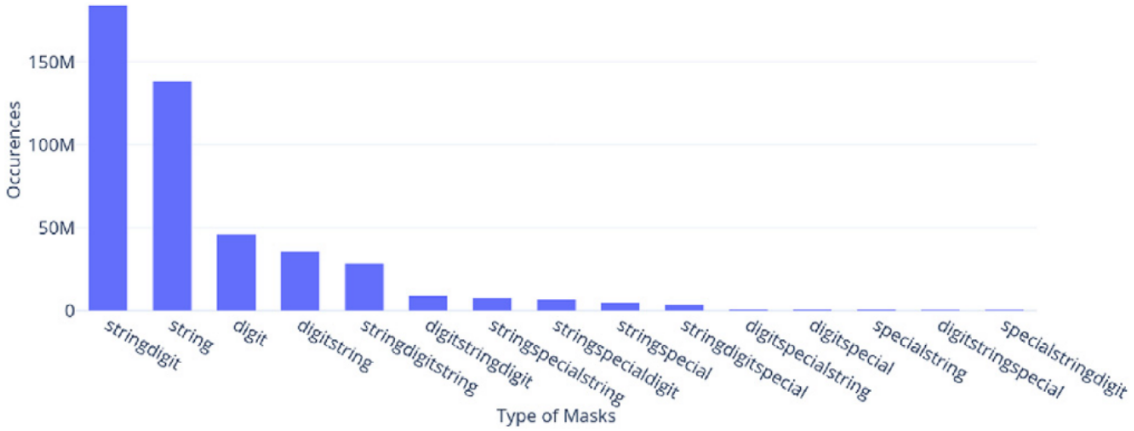
För avancerad analys används ett verktyg som kallas Oðinn Framework för att dela upp lösenord i grundläggande fragment för att hitta deras semantiska betydelse. Lösenorden



Figur 3.3: Lösenordskategorier

delas först upp i tre grundläggande teckenuppsättningar, nämligen bokstäver, siffror och specialtecken. Bokstäverna körs sedan genom en pythonimplementering av SymSpell, som är en av de mest effektiva stavningskorrigeringsalgoritmerna. Symspell kräver en datauppsättning av ord för att kunna fungera korrekt, och eftersom lösenord ofta använder slang och vanliga uttryck som inte finns i den typiska engelska ordboken, valde forskarna istället att använda en datauppsättning bestående av cirka 3.9 miljarder Reddit-kommentarer. Reddit-kommentarerna ger en mycket bra ordbok eftersom den innehåller slang och vanliga uttryck, och eftersom kommentarerna är skrivna på flera olika språk resulterar det i en flerspråkig ordbok.

Av de 3.9 miljarder icke-unika lösenorden från HIBP-datasettet producerade Oðinn omkring 1.575 miljarder fragment. Av dessa 1.5 miljarder var några av de vanligaste fragmenten bokstäverna äöch "i", samt korta ord som förekommer i det engelska språket som to", än", "inöch the". Lösenord som *qwerty*, *password* och *love* anses vara bland dom sämsta lösenorden och de förekom som förväntat i top 5 lösenordsfragmenten. Lösenordspromenader som qwerty, qwe, qaz, samt sekventiella siffror var frekventa i både ord- och



Figur 3.4: De 15 vanligaste lösernordsmaskarna

sifferfragmentslistorna. 3 av de första 5 sifferfragmenten var sekventiella siffor. Siffor, som 1, 2, 3, dubbelsiffror, som 12, 11, 13, och nummerrepetitioner, som 111111, 000000, uppkom frekvent i de top 50 sifferfragmenten. Då det kom till specialtecken var de 15 vanligaste specialteckensfragmenten enskilda. För fragment som inte bestod av singla specialtecken var det vanligaste mönstret upprepningar. Det är värt att notera att förekomsten av specialtecknensfragment var märkbart mindre än ordfragment och sifferfragment. Det kan bero på att användare undviker tecken som kräver flera tangenter att skriva, vilket ofta är fallet för specialtecken. Vid analysering av top 500 vanligaste fragmenten hittades frekventta nummerfragment inuti intervallet 1900-2020. Dessa kan antas vara vanliga årtal. Detta får forskare att tro att användare väljer lösenord på basen av minnesvärda mönster, som vikgita år eller datum. [8]

3.4 Hårdvarofaktorn

Det är viktigt att förstå hur snabbt lösenord kan knäckas med hjälp av hårdvara. Lösenord lagras oftast i ett hashat/saltat hashformat. Hashformatet är implementerat som en säkerhetsparameter ifall en dataläcka skulle ske. En snabb hashfunktion skulle möjliggöra att en attackerare kan testa flera lösenord än ifall hashfunktionen skulle vara långsam. MD5 hashfunktionen har använts i stor utstreckning för att lagra lösenord. Ett enda spelgrafikkort (Nvidia 2080 Ti) kan utverdera 50*10$\hat{9}$ olika lösenordskombinationer i sekunden. Grafikkortet skulle därmed lyckas testa alla 8 tecken långa MD5-hashade lösenord på ca två dagar. Vid användningen av en bättre hashing algoritm som t.ex BCRYPT som är designad för att göra beräkningarna långsammare nådde forskarna [8] hastigheter på ca. 25 000 ord per minut. Detta är vidare förstärkt då forskare lyckades nå hastigheter mellan 25 000 och 32 000 ord per minut då de försökte gissa lösenorden till några moderna kryptera-

de dokument, som t.ex ett låst Microsoft word dokument och en låst 7-zip fil. [2] Noterbart är dock att användningen av flera grafikkort skulle öka beräkningshastigheten linjärt samt att grafikkortens prestanda ökar med varje produktgeneration. Vid användning av ett Nvidia 2080 TI-grafikkort skulle ett lösenord som består av 15 siffror skulle knäckas inom en dag då lösenordet är hashat med MD5 och 1268.3 år då det är hashat med BCRYPT. För lösernord som består av 12 små bokstäver skulle det ta ca 22 dagar med MD5 och 120 961 år med BCRYPT. Om övrig information som tidigare lösenord eller personlig information är tillgänglig är det lättare att utföra specifika attacker. Ifall övrig information är given ökar sannolikheten att lösenordet knäcks, även för svårare lösenord.

//TODO add here about why this is important

3.4.1 Andra alternativ

//TODO list some alternatives for different applications

3.4.2 är människan fortfarande svagaste länken?

//Is it still the weakest link and what can we do to mitigate this? [11]

3.4.3 minimi antal försök

//TODO how does this effect the entropy level needed

Letter	Count	Number	Count	Special	Count
a	2.335%	1	8.240%	.	0.871%
i	1.168%	123456	5.137%	_	0.666%
qwerty	0.597%	123	2.574%	!	0.469%
password	0.510%	2	2.398%	@	0.334%
love	0.484%	123456789	2.083%	—	0.327%
my	0.356%	3	1.788%	:	0.140%
abc	0.274%	4	1.578%	#	0.105%
to	0.259%	5	1.111%	*	0.090%
an	0.259%	12	1.079%	$	0.071%
qwe	0.248%	7	1.029%		0.065%
in	0.238%	0	0.870%	&	0.045%
the	0.228%	8	0.812%	+	0.042%
qaz	0.223%	6	0.810%	?	0.037%
iloveyou	0.221%	12345	0.764%	,	0.035%
ws	0.217%	9	0.761%	/	0.031%
as	0.209%	1234	0.664%	!!	0.025%
no	0.198%	11	0.599%	.:	0.023%
ilove	0.196%	13	0.518%	&#	0.022%
by	0.191%	12345678	0.474%	=	0.021%
man	0.190%	01	0.430%	;	0.018%
baby	0.178%	10	0.425%	..	0.017%
on	0.176%	1234567890	0.418%	’	0.016%
it	0.156%	111111	0.411%	%	0.014%
we	0.145%	22	0.390%	<	0.014%
go	0.145%	23	0.375%	(	0.011%
he	0.145%	123123	0.365%	[	0.011%
asd	0.134%	1234567	0.360%	)	0.011%
sexy	0.131%	69	0.331%	**	0.010%
you	0.128%	21	0.321%	…	0.010%
boy	0.126%	14	0.284%	:,	0.009%
of	0.124%	15	0.248%	‘	0.009%
qa	0.117%	09	0.248%	$$	0.008%
girl	0.116%	08	0.236%	__	0.007%
fuckyou	0.114%	07	0.224%	!!!	0.007%
july	0.113%	99	0.224%	@@	0.006%
angel	0.111%	24	0.222%	–	0.005%
ma	0.109%	88	0.221%	.,	0.005%
march	0.107%	16	0.212%	^	0.005%
dog	0.106%	18	0.209%	~	0.004%
at	0.105%	000000	0.207%	!@	0.004%
big	0.103%	17	0.206%	! ~!	0.004%
monkey	0.102%	00	0.204%	>	0.004%
one	0.101%	19	0.202%	***	0.004%
alex	0.099%	77	0.193%	!@#	0.004%
red	0.095%	33	0.190%	]	0.003%
us	0.094%	20	0.187%	??	0.003%
qwer	0.094%	123321	0.183%	++	0.003%
qwertyuiop	0.094%	25	0.181%	”	0.003%
dragon	0.092%	666	0.174%	???	0.003%
life	0.091%	06	0.170%	==	0.002%
shark	0.090%	89	0.150%	*****	0.002%

Kapitel 4

Avslutning/resultat/takeaways etc?

4.1 Resultatet

Användare väljer lösenord som är enkla att komma ihåg och på grund av detta är de svaga och lätta för datorer att gissa. [8]

Då användare skapar allt flera lösenord och måste hålla reda på flera olika enheter blir används ofta samma eller liknande lösenord för att enklare komma ihåg dem.

Användarna använder gärna lösenord som är också korta så att de inte ska ta länge att skriva ut och för att då är de lättare att komma ihåg.

Reglerna om lösenord hjälper ofta att göra lösenordet starkare, men det hör också att lösenorden ofta ser liknande ut och innehåller liknande element, vilket gör dem enklare att gissa.

Användare skapar ofta lösenord relaterade till användarens omgivning, med användning av bekanta ord, namn eller saker i lösenordet, och det gör lösenordet enklare att gissa, speciellt ifall man har övrig information om personen.

//ADD A LOT HERE

Källförteckning

[1] URL: https://github.com/iphelix/pack.

[2] Oleg Afonin. *Breaking Passwords with NVIDIA RTX 3080 and 3090*. 2020. URL: https://blog.elcomsoft.com/2020/12/breaking-passwords-with-nvidia-rtx-3080-and-3090/.

[3] Mashael AlSabah, Gabriele Oligeri och Ryan Riley. ”Your culture is in your password: An analysis of a demographically-diverse password dataset”. I: *Computers Security* 77 (2018), s. 427–441. ISSN: 0167-4048. DOI: https://doi.org/10.1016/j.cose.2018.03.014. URL: https://www.sciencedirect.com/science/article/pii/S0167404818302979.

[4] Joe Dibbley. *What are NIST Password Guidelines?* 2023. URL: https://blog.netwrix.com/2022/11/14/nist-password-guidelines/.

[5] *Differences Defined and Which Is Better?* 2022. URL: https://www.okta.com/identity-101/password-vs-passphrase/.

[6] Lisa Eadicicco. *The Man Who Invented The Computer Password Admits That It's Become A Nightmare*. 2014. URL: https://www.businessinsider.in/the-man-who-invented-the-computer-password-admits-that-its-become-a-nightmare/articleshow/35484027.cms.

[7] *Have i been PWNED*. URL: https://haveibeenpwned.com/.

[8] Aikaterini Kanta m. fl. ”How viable is password cracking in digital forensic investigation? Analyzing the guessability of over 3.9 billion real-world accounts”. I: *Forensic Science International: Digital Investigation* 37 (2021). URL: https://www.sciencedirect.com/science/article/pii/S2666281721000949?via%3Dihub.

[9] Robert Morris och Ken Thompson. ”Password Security: A Case History”. I: *Commun. ACM* 22.11 (nov. 1979), s. 594–597. ISSN: 0001-0782. DOI: 10.1145/359168.359172. URL: https://dl.acm.org/doi/pdf/10.1145/359168.359172.

[10] Richard Shay m. fl. ”Correct horse battery staple: Exploring the usability of system-assigned passphrases”. I: *Proceedings of the eighth symposium on usable privacy*

and security. 2012, s. 1–20. URL: https://dl.acm.org/doi/pdf/10.1145/2335356.2335366.

[11] Viktor Taneski, Marjan Heričko och Boštjan Brumen. ”Password security — No change in 35 years?” I: *2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. 2014, s. 1360–1365. DOI: 10.1109/MIPRO.2014.6859779. URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6859779&tag=1.

[12] Dawn M. Turner. *Digital Authentication - the basics*. 2016. URL: https://www.cryptomathic.com/news-events/blog/digital-authentication-the-basics.

[13] Ding Wang m. fl. ”Zipf’s Law in Passwords”. I: *IEEE Transactions on Information Forensics and Security* 12.11 (2017), s. 2776–2791. DOI: 10.1109/TIFS.2017.2721359. URL: https://ieeexplore.ieee.org/document/7961213.

[14] Shannon Williams. *Average person has 100 passwords - study*. 2020. URL: https://securitybrief.co.nz/story/average-person-has-100-passwords-study.

[15] *xkcd passphrase comic*. URL: https://xkcd.com/936/.