



Batman Hacked My Password: A Subtitle-Based Analysis of Password Depiction in Movies

Maike M. Raphael, Leibniz University Hannover; Aikaterini Kanta, University of Portsmouth; Rico Seebonn and Markus Dürmuth, Leibniz University Hannover; Camille Cobb, University of Illinois Urbana-Champaign

<https://www.usenix.org/conference/soups2024/presentation/raphael>

**This paper is included in the Proceedings of the
Twentieth Symposium on Usable Privacy and Security.**


August 12–13, 2024 • Philadelphia, PA, USA

978-1-939133-42-7

**Open access to the Proceedings
of the Twentieth Symposium
on Usable Privacy and Security
is sponsored by USENIX.**

Batman Hacked My Password: A Subtitle-Based Analysis of Password Depiction in Movies

Maike M. Raphael 
Leibniz University Hannover

Aikaterini Kanta 
University of Portsmouth

Rico Seebonn 
Leibniz University Hannover

Markus Dürmuth 
Leibniz University Hannover

Camille Cobb 
University of Illinois Urbana-Champaign

Abstract

Password security is and will likely remain an issue that non-experts have to deal with. It is therefore important that they understand the criteria of secure passwords and the characteristics of good password behavior. Related literature indicates that people often acquire knowledge from media such as movies, which influences their perceptions about cybersecurity including their mindset about passwords. We contribute a novel approach based on subtitles and an analysis of the depiction of passwords and password behavior in movies. We scanned subtitles of 97,709 movies from 1960 to 2022 for password appearance and analyzed resulting scenes from 2,851 movies using mixed methods to show what people could learn from watching movies. Selected films were viewed for an in-depth analysis.

Among other things, we find that passwords are often portrayed as weak and easy to guess, but there are different contexts of use with very strong passwords. Password hacking is frequently depicted as unrealistically powerful, potentially leading to a sense of helplessness and futility of security efforts. In contrast, password guessing is shown as quite realistic and with a lower (but still overestimated) success rate. There appears to be a lack of best practices as password managers and multi-factor authentication are practically non-existent.

1 Introduction

Cybersecurity is a topic that virtually everyone encounters every day, from the first unlocking of the smartphone in the

morning to reading emails at work or communicating with friends at night. This requires many decisions about which links to click, which websites to trust or which password to choose. Among other things, these decisions are based on knowledge and beliefs about the subject area in question that determine, for example, what is perceived as “secure” or “insecure” [20, 73]. It is therefore important that this knowledge is correct and beliefs are aligned with reality. However, studies show that for cybersecurity these are often incorrect or incomplete, leading to “bad” security practices [1, 63].

This problem becomes particularly evident in password security, which is an area where many misconceptions are found. Various studies show that people do not know the characteristics of good passwords, do not know how to handle passwords in general or do not remember the recommendation to change the password regularly, which has been proven to be bad advice in recent years [11, 24, 45]. This is a big issue because, despite their weaknesses and numerous alternatives being available, passwords are still by far the most common form of online authentication [26]. It is therefore important that the understanding of password security and good password behavior is reinforced.

One source that influences people’s perception of cybersecurity is likely films [54, 73]. Literature shows that people learn from media and use it as a source of information [51, 54]. Films play a major role in this; they have been a popular entertainment medium for decades and are watched by thousands of people every day. Therefore it is hardly surprising that films influence knowledge and behavior [19] or the attitude toward technology [9] and that this may change how people handle specific topics and make decisions. Because people often cannot decide if what they see is realistic or not, they are in danger of taking fictional portrayals as realistic which may influence their thinking about certain topics including cybersecurity [20, 73]. It is therefore important to ensure that things are presented in a good and realistic way so that people potentially learn something *true* from them [12].

The use of certain technologies in movies “both reflects and influences society’s use and attitudes toward the portrayed

Copyright is held by the author/owner. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee.

USENIX Symposium on Usable Privacy and Security (SOUPS) 2024,
August 11–13, 2024, Philadelphia, PA, United States.

technology” [9]. Literature shows that what we see in the media (partially) reflects life [19] and shows what our society thinks and is interested in [21, 67]. Furthermore, movies can be regarded as “Cultural Artifacts” and historical snapshots” [9] enabling us to compare attitudes from different years. So we can use those to learn something about our society during the ages and to identify password behavior that seems to be considered typical as is already shown impressively for society’s attitude toward technology.

Regarding the many misconceptions concerning passwords, it is to be expected that their portrayal in films may reflect an outdated or insecure password behavior that we also see in society. At the same time, showing good passwords and secure behavior could have a significant impact on the understanding and the overall security of people regarding password usage. Consequently, the aim of this paper is to answer the question of how accurately cybersecurity topics in general and password-related topics in special are depicted in films.

In this paper, there are four primary contributions to improve understanding of password depiction:

- A subtitle analysis as a novel approach to analyzing the occurrence of passwords within a large amount of data: 97,709 movies of various genres, from 1960 to 2022.
- A collection of which films and scenes and in which context passwords play a role and a statistic evaluation of the results from 2,851 movies. This includes what the password is used for, different kinds of password behavior and (missing) best practices such as the use of password managers and multi-factor authentication.
- An evaluation of the strength of the passwords shown in films and a linking of this knowledge with the results from step two to understand the role of *strong* passwords in movies.
- In particular, an investigation is conducted on attacks on passwords to understand whether passwords are presented as “secure” and which circumstances lead to blighting password protection. This includes watching selected movies to understand the overall importance and ambient conditions of password attack movie scenes.

The results show how many everyday password activities are mirrored in films and how often the topic appears in a wide variety of genres and years. However, this often involves insecure behavior such as careless sharing by writing down or reusing passwords for different accounts. Good password practices such as using password managers or multi-factor authentication is scarcely depicted in films. Even if strong passwords are used, these hardly increase security – passwords within the highest strength category (as shown in Section 6) are guessed as often and easily as very bad ones. Both hacking and guessing attacks are often frighteningly successful, which gives the impression that passwords can hardly withstand any attacks. However, there is a strong contrast between the very unrealistic hacking attacks and password guessing, which is often portrayed in a very realistic manner.

2 Background & Related Work

We first describe a small but closely-related body of work focused on movies and cybersecurity. We looked to the fields of film studies and linguistics to inform our understanding of methodological best practices and the way that movies impact people and society. Finally, since we focus on the depiction of passwords, existing knowledge about passwords, password security, and user experiences with passwords provided important context for structuring our analysis and interpreting our findings.

Cybersecurity and Movies Prior work has shown that many users learn about cybersecurity from media, including advertisements, news, and fictional narratives such as television and movies [51, 52, 54].

Specifically, researchers have studied the impact of movies on people’s understanding, perspectives, and behaviors related to hacking [4, 20], biometric and non-biometric authentication methods [73], and technology broadly [9], finding that movies have the capacity to misinform people or guide them toward a better understanding of technology. Authors argued that the movies sometimes confirmed participants’ existing mental models, for example beliefs that only famous or rich people will be attacked and that – if targeted – security measures are futile anyway [20]. Other prior work helps us understand the mechanisms through which media such as movies might influence people [30]. For example, mental models are thought to be an important aspect of decision-making [30]; thus, researchers finding that watching movie scenes impacted mental models [4, 20] suggests that movies could influence people’s decisions and behaviors as well. Perhaps the visual and/or video format of movies contributes to their ability to influence people; studying the difference in impact of a video-based message or a text-based message, Albayram et al. found that people who watched videos were more likely to adopt password managers [2]. Another influential factor may be the narrative structure that is common in fictional media such as movies. Prior work has repeatedly found that we learn about security through stories [46, 48].

Since movies have the capacity to (mis)inform, it is pertinent to understand their contents and to what extent this content is realistic. Examples poking fun at the inaccuracies of cybersecurity in movies are easy to find in blog articles [53], online repositories of TV tropes [16, 17], and even in a talk at DefCon [38]. These sources emphasize inaccuracies such as hacking or decryption being absurdly easy or quick, technical terms being thrown around without real meaning, and images of illuminated screens with rushing lines of code. Similarly, Christmann et al. find inaccuracies with password advice in YouTube videos and propose a list of requirements for security awareness videos dealing with password behavior [12]. In a more systematic study, Gordon assembled and analyzed a data set of 50 “hacker movies” from the 1960s

through the early 2000s, comparing the key themes in these movies with reality [23]. Gordon found that some aspects of movies' portrayals of hackers was quite realistic (e.g., finding that the inaccurate "stereotypical view of outsider attacks by teenagers" is *not* coming from this set of movies), while some were not (e.g., the ratio of insider to outsider attacks), but argues overall that these movies are likely to be a useful resource for security course instructors.

Learning From and With Film Media Film media (i.e., television and movies) can have a positive influence on adults' or children's learning and influence them to adopt beneficial attitudes and behaviors [18, 34, 35, 65, 68, 69]. For example, Whittier et al. deployed an online survey shortly after a popular television show had aired an episode with a story line about syphilis and found that participants who had seen the episode reported higher intention to be screened for syphilis [68]. In contrast, movies can also have harmful learning effects. For example, misrepresented medical scenes can lead to dangerous misconceptions that reinforce racist stereotypes [44] or lead to self-diagnosing with insufficient medical understanding [49]. This has led to the creation of programs for reviewing movie contents [47, 60]. Hoffman et al. found that medical television's influence on viewers' health-related knowledge was deemed negative in 11% of prior studies, positive in 32% and mixed in 58% [27].

Unlike the formats that we typically associate with the idea of "learning," people learn *passively* from film media [35]. Krugman and Hartley assert that this type of passive learning is "characterized by an absence of resistance to what is learned" and so in some ways has capacity to be especially powerful [35]. But the precise impact a movie on a particular person is likely unpredictable. Fearing argues that "what the individual 'gets' [from the movie] is determined by his background *and his needs*. He takes from the picture what is usable for him or what will function in his life" [19]. Integrating movies as a tool for active teaching/learning has been widely discussed in fields such as medicine [6, 36, 65], counseling [34], and international politics [18].

Film studies is a rich field whose methods often involve close watching of one or a small set of films, sometimes frame-by-frame analysis, and factoring in how elements such as the film creators' personal backgrounds and societal or cultural issues that may have influenced the film itself and its reception by audiences [13, 56]. The increasing availability of digital analyses, which enables "big data" in film studies, has shifted approaches and spurred, for example, the establishment of the Digital Cinema Studies network [57]. Subtitle analysis has been used to gain insights about the contents of bigger sets of movies. For example, linguistics researchers studied speech acts by analyzing "Evim Sensin" (You Are My Home, 2012) subtitles [29] and compared word frequencies in Greek and Polish movie corpora [15, 40]. These analyses found that language in subtitles is similar to everyday language and

that topics from society are reflected in films. Other research is based on searching for words in subtitles to analyze, for example, hate speech or physical aggression and verbal insults within selected movies [61, 70] or how sex behavior is referenced in a Netflix series [71].

User-Focused Password Research There is a significant body of Usable Security & Privacy research regarding passwords and technology users, which seeks to answer questions such as: What are users' existing password practices [33, 41, 58, 62]? What do users understand or believe about passwords, password strength, and password attacks [33, 42, 58]? In what ways are passwords typically attacked [50]? How can we encourage users to create better passwords or otherwise decrease vulnerability to authentication attacks [2, 55, 72]? Common practices that make users' accounts more vulnerable include creating predictable passwords [58], re-using passwords across different services [41], and using personal information in a password (e.g., year of birth, names of relatives) [62]. Users also regularly expose these types of personal information online [28]. Analyses of leaked passwords show that users often add numbers at the end of their passwords, capitalize the first letter of the password, and make common letter substitutions (e.g., "@" to replace "a", or "1" to replace "i") [33]. Users tend to overestimate the security of passwords they create [58] and have different misconceptions regarding password composition, handling and attacks [42]. Security researchers have also formed an understanding of how passwords (or authentication systems more generally) are typically compromised. This can happen via automated password guessing (e.g., brute force or dictionary attacks) or compromising other parts of a user's security (e.g., deploying a keylogger or using social engineering to get a user to reveal their password) [50].

3 Method

In previous work so far only a targeted selection of films have been examined and with a very specific focus on hacking, so we take an approach that enables us to draw quantitative conclusions about passwords in movies. To scale our analysis via automation, we used text-based approaches to analyze a large set of movie subtitles.

Creating a Movie Subtitles Dataset We obtained the subtitles from a torrent link posted on the social news aggregation website Reddit *r/DataHoarder* [3]. The torrent contains a database (`opensubs.db`, 136.8 GB) of 5,719,123 subtitle files, crawled on July 24, 2022. It also contains a metadata file (`subtitles_all.txt.gz`, 309 MB) that includes information such as movie name, year, language, content type (movie, TV show), season, episode, IDs (IMDB, OpenSubtitle), upload date, frame rate, and file format.

The Reddit thread stated that these subtitles were initially sourced from the website [opensubtitles.org](https://www.opensubtitles.org) [8], one of the largest subtitle databases on the Internet. Subtitles are uploaded by users, who then vote and comment on the quality of subtitle files.

We filtered out non-English subtitles and subtitles for content besides movies because the full torrent also contained subtitles for TV shows and other content types. Movies that appeared in a non-English language (e.g., *Parasite*, 2019; Korean) but had English subtitles available were included in the analysis. Additionally, the torrent contained duplicate subtitle entries for some movies (e.g., if two users had uploaded subtitles for the same movie); we removed duplicates by always using the most recently-added subtitle file. Our final dataset contained subtitles and metadata for 97,709 movies. More information about this movie subtitle dataset (e.g., graphs of their genres and years of distribution) can be found online¹. We obtained additional metadata including genre, popularity, and other details from *The Movie Database* (TMDB), using the TMDB API.

Identifying Password-Related Content in Movies To automatically identify content in movies that is related to our research topic, we perform a search within the subtitles for the word *password*. We found that this straightforward approach was the most appropriate for identifying relevant content in such a large dataset. We considered including other authentication-related words or phrases in our keyword search, including *passphrase* (occurs only seven times in the dataset) and *PIN* (high false positive rate due to semantic overload). *Password* appeared 5,982 times in 2,851 different movies (just under 3% of movies in our dataset).

To create units of analysis corresponding approximately to the notion of a movie “scene,” we considered the nine subtitle lines before and after the occurrence of the word *password* (i.e., a total of 19 lines). Note that subtitles do not encode the idea of a “scene;” we found this to be a conservative approximation (i.e., the researchers agreed that 9 lines before/after the keyword were more than enough context to meaningfully interpret the data). Subtitles include newlines corresponding to what would appear as one line of text on someone’s screen if they were watching the movie with captions. There is not information about who said which words. A longer dialogue from one character may span multiple lines. Typically (but not always), newlines or other visual indicators such as dashes are inserted when a new character begins speaking. Subtitles typically contain (most of) the spoken words, though cross-talk (i.e., multiple people speaking at once) and background dialogue may not be fully captured. Sometimes subtitles contain additional information about the audio such as “laughter” or “music”. In the results, we report on patterns of how these

¹<https://www.itsec.uni-hannover.de/de/usec/forschung/medien/password-depiction-in-movies>

movies with password-related content are distributed in terms of year and genre.

Characterizing Scenes about Passwords We started with an open-coding approach to analyzing these 5,982 password-related scenes. Two authors each used MaxQDA to independently open code the same set of 50 scenes, which included 10 randomly selected scenes from each of five time intervals (including very old and very recent movies). The researchers then compared their open codes and generated a codebook. One author applied the codebook to all scenes. When coding decisions were unclear, he consulted with co-authors to reach a consensus. The codebook can be found online¹ alongside with a list of all scenes including the movie metadata and the set of codes we applied to each scene.

Analyzing Password Topics and Password Attacks We characterized the context of use for the password (e.g., if the password is used for a computer, a website account or locks), and different activities that are performed with passwords (e.g., password creation, change or losing a password). In particular, we coded the scenes based on whether they contain *password hacking* and/or *password guessing*. These codes were used as the basis for generating a sample of movies that we watched manually. We report summary statistics and patterns that emerge between these codes and also over time/by genre, and we include relevant quotes from the subtitles to illustrate our findings and provide qualitative depth.

Measuring Movie Passwords’ Strength While applying the codebook, we recorded all passwords that showed up in the transcripts (e.g., “*Your password’s 999999?*” (Max Winslow and *The House of Secrets*, 2020)). In cases where passwords were described verbally, we recorded our best approximation of the plaintext passwords (e.g. in (*The Disappearance of Jennifer Dulos*, 2021), subtitles state “*What’s the password? - It’s four zeros*”; we recorded this as 0000). This resulted in a list of 687 passwords which are listed in Appendix A organized by strength using the *zxcvbn* metric as described below. To measure the strength of these passwords, we applied two well-known password metrics:

- *zxcvbn*. A simple but relatively accurate [22] strength meter developed by Dropbox [66]. *zxcvbn* categorizes passwords into 5 strength categories from weakest (Class 0) to strongest (Class 4). Passwords up to Class 2 are easily guessable, whereas Class 4 contains passwords that would require 10^{10} guesses [33].
- *PGS*. The “Password Guessability Service” created by researchers at Carnegie Mellon University [59]. The output of this metric is the number of guesses it would take to guess the password (or -5 if it cannot be guessed).

We compared the strength of these passwords with real-world leaked passwords from several well-known breaches:

- The Popular-200 (200 most used passwords of 2023) [43], a current dataset with a focus on frequently used (and therefore tending to be less secure) passwords.
- The *Ignis IM* wordlist [25], which was assembled in 2020 from various data leaks (Collection #1, Dropbox, LinkedIn, and others) and contains the 1 million most popular passwords found within those data leaks [32].
- A list from the *RockYou* data breach [14], which was leaked in 2009, but contains the full distribution of passwords from the weakest to the strongest since it was leaked in plaintext and is frequently used in comparable research.

Watching Movies to Gain Deeper Qualitative Insights To find additional information that could not be found out by analyzing only a specific scene (such as the importance of the password activity for the whole movie) or could not get out of the subtitles (such as a password which is typed in but never said out loud, which is why it may not appear in the subtitles) and to compare the findings of our subtitle analysis with some real movie scenes, we watched a small subset of 21 movies and evaluated the password attack scene in the context of the overall movie plot. Considering movies that came out between 2013 and 2022 which were sorted from most to least popular based on the total amount of votes a movie reached on TMDb [5], we included the top 10 movies that contain *password hacking* and the top 15 movies that contain *password guessing*. Three movies were in both categories (i.e., contained hacking *and* guessing scenes), and one movie was not available to watch online.

Six people participated in this task, watched the allocated movies in full, and filled out a questionnaire that was discussed and iterated on by the authors; the questionnaire aimed to capture details that would have been missed in transcript analysis. All six people who participated in this task were trained and had opportunities to ask questions about their understanding of the questionnaire before starting. The questionnaire and list of watched movies can be found online¹.

High-Grossing Movies Our dataset includes both very popular and relatively obscure movies. Popular movies have (by definition) already reached a broader audience and will likely continue to be viewed more often, which makes their capacity to (mis)inform viewers especially important to consider. While our analysis is primarily concerned with the full dataset, we assembled a secondary High-Grossing dataset to assess whether high-grossing movies' characteristics are meaningfully different. The High-Grossing dataset consists of all movies in our dataset that appear in the list *Top 1000 Highest-Grossing Movies of all Time* [7] as of April 2024. This list contains 39 movies that are newer than our dataset and three that are not available in English (i.e., excluded from our dataset). Of the remaining 958 movies, 70 (7%) contain

the word *password* at least once (listed in Appendix C). We compare the High-Grossing movies with our other findings in Section 7.

Ethical Considerations Movies are subject to copyright; however, analysis like ours should be protected under fair use. Additionally, the user-generated subtitles on OpenSubtitles.org seem to not infringe on copyright [64]. While OpenSubtitles.org disallows scraping, they explicitly allowed non-registered users to download subtitles at the time of database creation and still allows non-commercial, scientific, and educational use [64]. An older corpora from this site was published in 2018 and is used by the linguistics community [37]. By using a dataset that had already been scraped and shared publicly, we avoided stressing OpenSubtitles.org's server bandwidth. Finally, while the subtitles are user-generated, we did not use data about any of the individuals who uploaded them.

Limitations Some of our methodological choices present limitations to how readers should interpret our results and what we could find. However, these trade-offs enabled for a much broader analysis than has been conducted previously, and so represent a deliberate choice. Searching for only the keyword *password* almost certainly excluded relevant scenes and movies (both content that is relevant to passwords specifically, but also content more broadly related to authentication or security and privacy in general). Using only subtitles limits what information we can analyze, and we must assume that the subtitles sometimes leave out relevant context or are misleading. Our deep qualitative analysis through watching a small number of movies helps address this concern (see Section 5). It is possible that we could have understood the scenes marginally better by including more than 19 subtitle lines in our analysis, though we found this to be acceptable empirically. The subtitles used in this work were user-contributed (on opensubtitles.org [8]) and some of the subtitles were translated, which may change the meaning of individual sentences. We only informally checked the data quality (by paying attention to subtitles during the movie watching activity), but we found that the subtitles were highly accurate. Finally, while this paper is the largest study on how cybersecurity-related topics are presented in entertainment media, we do not consider other types of content besides movies, including TV shows (or series) or other online media.

Additionally, it must be emphasized that our analysis focuses exclusively on the depiction of password topics in films and thus on the question of what people are shown when they watch the movies - and could potentially learn from them. Whether and to what extent people actually do take away this information from films remains to be investigated. Therefore, there is no evidence presented that these scenes have any real-world impact.

4 Results: Depiction of Passwords in Movies

Next, we present the results of our analysis. In the first step, we will look at which movies contain password-related content at all, before turning to different scenes and what is done with passwords in them. This includes which things passwords are used for, which activities are described in the context of the passwords, and more.

4.1 Movies Featuring Passwords

Recall that only around 3% of movies in our subtitle data set contain the word *password* (2,851 movies). We start by asking about the characteristics of these movies. Are passwords more likely to be mentioned in certain genres of movies? How has the frequency of mentioning passwords changed over time?

Figure 1 shows how movies whose transcripts include *password* are distributed across various genres are more likely to contain *password*. *Password* is mentioned most commonly in Thriller, Science Fiction, and Action movies and least frequently in Western, Music, and Documentary movies.

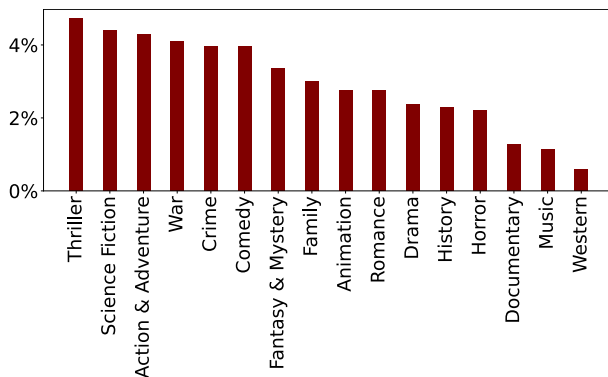


Figure 1: Percent of movies within each genre that contain the word *password* at least once. Thriller is the genre that mentions passwords most often, with 5%.

As shown in Figure 2, the proportion of movies containing *password* increased over time which means that newer movies are more likely to mention *password* than older ones. For movies that have come out since the start of 2020, slightly over 5% (or 1 in 20 movies) contain the word *password*. 71% of movies with *password* have been released since 2005.

Comparing the High-Grossing movies with the overall dataset, a comparatively high number contain password scenes (7%, compared to 3% in the overall dataset or 5% of all movies since 2020). This may be explained by the makeup of the High-Grossing dataset: they tend to be newer (545 came out after 2010) and they tend to fall into genres that we found more often reference passwords (78% are Thriller, Sci-Fi, or Action & Adventure). We return to this comparison in 7.

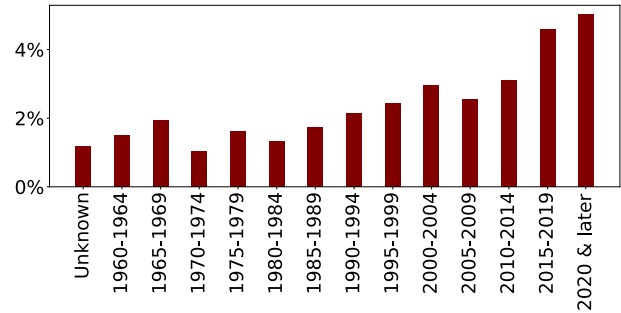


Figure 2: Indication of what percent of movies within a year interval contains the word *password* at least once. In recent years, slightly more than 5% of films contain the word *password*, in past years this was between 1% and 2%.

4.2 Password Behavior in Movie Scenes

Next, we present the types of password behavior that we detected within the analysed movie scenes as well as patterns that emerge based on applying these codes to the data.

Context of Use We were often able to use the subtitles to discern what the context of use was for the password referenced in a particular scene. We categorize these contexts into computer-related, Internet-related, banking, (interpersonal) legitimation, or anything else, as shown in Figure 3. A more detailed breakdown of contexts of use is contained online¹; here we primarily report findings related to these high-level categories. *Interpersonal legitimation* contains all scenes in which a person uses passwords or passphrases to prove towards other people that they belong to a specific group or are allowed to perform a certain activity (e.g., enter a restricted area, perform an operation as a spy, etc.). Example scenes are the following: “*It had to be a Gryffindor. Nobody else knows our password*” (Harry Potter and the Chamber of Secrets, 2002); “*you’ll get your instructions day by day. The password for the contact will be »Wee-wee Birdie«, and the contact will answer »Poo-poo birdie«* (Brigada explosiva: Misión pirata, 2008). The most common *Internet-related* passwords are Wi-Fi passwords (including passwords “*to the internet*” (Witness to Murder, 2019)) and passwords for website accounts such as the “*registration in a site for dating*” (Love.net, 2011), or for the “*website you have an administrator’s account, right?*” (Suicide Club, 2018). *Other* combines a collection of all scenes that do not fit into the other categories. This includes, for example, when a password is used as a signal to start a specific activity as in “*At the password, »The cat is in the kitchen cupboard«, you’ll open the envelope*” (In Danger and Dire Distress the Middle of the Road Leads to Death, 1974). Other examples in this category are when a password is used to control the people’s minds or (in one case) pets, or when it is used as a mantra such as “*and remember the password: relaxation*” (The Big Bluff, 1995).

We observe shifts in the most frequent context of use over time; until 1990-1994, the dominant context of use was legitimation. In more recent time intervals, digital passwords used in computer contexts have become most common, and from the 2015s onward, Internet-related passwords have been dominant.

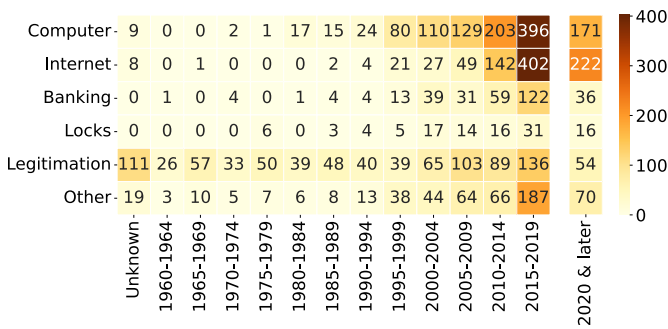


Figure 3: This figure shows how the context of use varies over time. Darker cells indicate a larger number of scenes with this context of use and release date, and the number in each cell conveys the number of scenes. Notice that legitimation was more common than other contexts in older movies, but computer- and Internet-related contexts have become most common recently.

Password Life Cycle In some scenes, characters speak about specific points in a password life cycle. Our coding process distinguished: password creation, changing a password, training to remember a password, losing or forgetting a password, performing a password recovery or reset, and reusing a password for different accounts. The distribution of these in scenes over time is shown in Figure 4. More detailed descriptions and examples of each point in the life cycle are included in Appendix B. Except for password recovery and password reuse, all codes appear with similar frequency in the movies; they are named between 104 and 115 times.

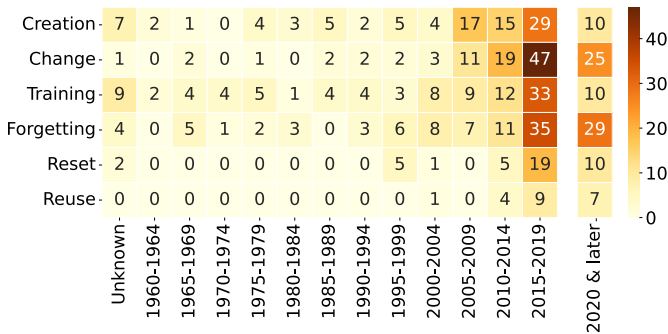


Figure 4: Points in the password life cycle and the number of scenes they occur in over time. Darker cells correspond to a larger number of scenes.

Password Sharing A password is shared in scenes from more than 1000 movies. Here, we consider *how* this happens and *with whom* the password is shared.

Intended/intentional sharing is the most common type of password sharing, presented in 469 scenes. Of those, 74 scenes include a deeper explanation of how the password should be handled, such as, “*Once you pass through the first step, you will receive a password on your phone. The last step is the key.*” (Collectors, 2020). However, password sharing is often non-consensual or forced (132 scenes). For example, “*The silver bowl your brother-in-law got from Turkey... Do you know how much its worth? Do us a favor, Just give the password to the lock of all the precious things in the house*” (French Biriyani, 2020). In 64 scenes, a password is shared unintentionally, such as in the following scene: “*You shouldn’t leave shit lying about. -How’d you get the password? -You had it taped underneath the fucking thing.*” (Boy A, 2007).

In 367 scenes, the password is distributed to one legitimated person such as a friend, colleague or family member. In most of these cases, the recipient takes on the role of a friend (186 scenes), followed by family members (78 scenes), work and the partner. In 93 scenes, it is shared with a small group of people, e.g., some direct colleagues or people from the same squad. In nine scenes, a large group is the recipient. For example, a password is forwarded via radio to all military units or “*the whole FBI*”(Enemies of the State, 2020).

Security Best Practices As described in the background section there is a large amount of security advice regarding password behavior. In addition to general recommendations such as not sharing passwords (discussed above), using password managers and multi-factor authentication are recommended as specific best practices. We have therefore analyzed what role they play in movies.

Password managers appear in only four scenes, and all four scenes are in one movie. It is the French movie *Disappear: Cover your online tracks* from 2021, a documentary including (among several other topics) a description of what a password manager is and how it can be used.

In seven movies – all from 2013 or later – the password is combined with a one-time password (OTP). That is, Multi-Factor Authentication is shown in these movies. In four scenes, the OTP is sent to a smartphone, in one to a key card, and in one it is created by “*a pair of watches that had undergone a special magnetization process. Only by putting the two watches together will a person be able to acquire the correct account and one-time password, thus, gaining access to the huge sum*” (Arjun Suravaram, 2019). In four scenes, the OTP is used to access a bank account or transfer money; in one, it is used for accessing a high-security area in a building, and in another, it is used to perform a password reset via phone (the OTP is sent to the phone and has to be read aloud).

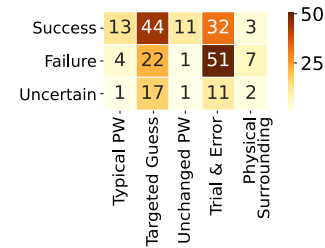
5 Depiction of Password Attacks

In the coding process, we distinguished two basic kinds of password attacks: *Password guessing*, where a human actively guesses candidate password based on frequent passwords, specific knowledge about a person, or known old passwords, and *password hacking*, where other techniques are used to obtain the password such as social engineering or shoulder surfing or using automated tools for (brute-force) guessing. Password guessing appears in 220 scenes and password hacking in 63.

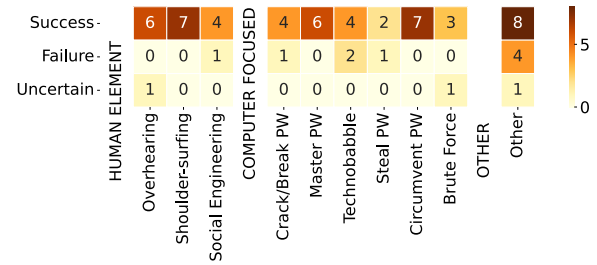
Password Guessing We observed different guessing approaches: Sometimes people try *typical passwords*, hoping that the target chose one of the easily guessable ones, e.g. “tell him that 1-2-3-4 as a password is worth fuck-all” (Todos tus secretos, 2014). Others use their knowledge about the person using the password and try to guess it by characteristics of the target, e.g. “The name of her boyfriend is Troy. But she calls him Batchoy. Her birthday is July 6th. So let’s try this” (Walwal, 2018). Sometimes, attackers know old passwords and hope they have not been changed: “Paul’s used the same combinations and passwords since we were freshmen at Cornell” (Consensus Reality, 2018). In some cases, the attacker looks around the physical surroundings and searches for (physical) clues or written-down passwords, for example “she scoured their apartment looking for passwords to get into his laptop” (Trust No One: The Hunt for the Crypto King, 2022). In other scenes, the attacker just tries whatever passwords come to mind. Most frequently, people in the movies use *trial & error* (94 scenes, 43% of all scenes with password guessing), closely followed by using knowledge about the person (83 scenes, 38%).

Overall, the guessing is successful in 103 scenes (47% of scenes with password guessing), 85 attempts fail (39%) and 32 (15%) have an uncertain outcome. Figure 5(a) shows guessing attempt success broken down by approach. Attacks based on knowing an unchanged password have the best success rate (11 out of 13, 92%). Using “typical” passwords and targeted guesses had a higher-than-average success rate (76% and 53%). Trial & error and using the physical surroundings have higher than average *failure* rates (54% and 58%, respectively).

Password Hacking We define *password hacking* to include all attack techniques except *guessing* (e.g., social engineering, shoulder surfing, overhearing passwords, using malware, virus software, keyloggers, or password cracking tools). In applying our codebook to the subtitles, we categorized the 63 scenes with hacking attacks into: (1) 19 scenes that focus on the human element (i.e., overhearing, shoulder surfing, and social engineering), which occurs in around 30% of these scenes; (2) 31 scenes that focus on the computer (i.e., breaking passwords, using master passwords to circumvent individual passwords, virus, brute-force, etc.), in almost 50%



(a) Guessing Attacks



(b) Hacking Attacks

Figure 5: Showcasing the amount of cases where codes applied to (a) guessing or (b) hacking and authentication outcomes overlap thus highlighting which attack strategies lead to which outcome of authentication. Cells are intensifying in shades of red proportionally based on the prevalent guessing strategies authentication outcome.

of hacking scenes, and (3) other in 21% of scenes. Hackers often performed some kind of “computer magic,” described with *technobabble* in the movie dialogue, as in the following scene from (America: The Motion Picture, 2021): “After a reverse hash, I backdoored the root password. A base checksum against the main data store allowed me to retrieve the salted hash, and then, from there, I was gleaming the cube.”

Password hacking in movie scenes tends to be even more successful than password guessing (81% success, 14% failure, 5% uncertain; see Figure 5(b)). Human-factor attacks are successful in 89% of scenes (17 of 19 successful attacks).

Deeper Insights from Watching 21 Movies with Hacking and Guessing

Here we convey our findings from watching 21 full movies that contain password hacking and/or guessing. We examine three key details that we were unable to analyze based on subtitles alone: the roles and character traits of attackers and their targets, more nuanced understandings of the assets targeted in password attacks, and the importance of the attack to the overall plot of the movie. We also consider the extent to which our subtitle analysis may have been incorrect or incomplete within the factors included in our codebook.

In movies, there is generally a clear distinction between “good guys” and “bad guys,” main and secondary characters, and it is generally easy to understand aspects of character development such as character traits (e.g., computer exper-

tise) and relationships between characters. While it was too difficult to discern these roles and character traits based on subtitles alone, we now paid special attention to these. Out of the 21 movies we watched, the character trying to hack or guess a password was only “bad” in two movies (i.e., in 90%, the person doing the hacking was good or neutral). In movies with hacking, the character(s) performing the attack were always main characters (if hacking occurred in a team, at least one team member was a main character); the characters trying to guess passwords were mostly of average importance (e.g., a friend or family member of the main character). Characters (or at least one team member) performing hacking had high computer knowledge, but those trying to guess passwords mostly did not have high computer expertise. These characters included: two superheroes, one police investigator (assisted by Batman), and a couple of gangsters, one of whom is a hacking specialist. The targets of password hacking attempts were typically opponents (good or bad) of the attacker, with no close relation except being rivals. Password-guessing targets were only active opponents in two cases, and they included family members (5 movies), characters with a romantic background (2 movies), colleagues or friends (2 movies), and neighbors (1 movie).

By watching movies, we were also able to better understand assets targeted by guessing or hacking attacks. Subtitles often mentioned only *the computer*, but we could not assess what role *this computer* plays in the plot. Within the movies we watched, we observed that hacking tended to target civic or company assets. For example, civic assets include targets in the context of secret agents and similar, which are of great (civic) importance such as the computer system of S.H.I.E.L.D.,² secret online videos concealed by criminals or a city’s traffic control system. Company assets that were targeted include the code to the space station, company servers with precious software, or employee data. In contrast, password guessing tended to target private assets, and the targeted assets themselves were of little importance to the plot (e.g., private computers and smartphones, two email accounts and once “all my private accounts”). Password hacking tended to be of high importance for the story (in all but one movie with hacking). For example, successful hacking saved thousands of lives in *Captain America: The Winter Soldier* (2014), circumvented the next crime in *The Batman* (2022), and resolved the entire plot of *Focus* (2015), a movie in which getting the secret code constitutes the main story line. Password guessing was mostly less important to the plot. For example, in *Blended* (2014) the main character’s son finds out she was on a date, but he would have found out a bit later anyway. We rated the password guessing as having “medium” importance for only two movies; in both cases, the guessing is only one of several steps to reach the final goal (e.g., freeing the second

²“Supreme Headquarters International Espionage Law-enforcement Division,” a fictional counter-terrorism intelligence agency from the Marvel cinematic universe (we watched *Captain America: The Winter Soldier*, 2014).

main character, who then helps to finalize the next quest in *Ready Player One* (2018)).

Finally, we specifically sought a deeper understanding of whether our analysis of subtitles alone was misleading or incomplete within the set of topics included in our codebook. Overall, we found little evidence that our subtitle analysis was insufficient. The six specific passwords that appeared in movies we watched were also present in the subtitles, though in some cases, the password appeared on screen before it appeared in the subtitles. For password guessing attempts, no tools or computer activity was ever shown; instead, in most cases characters inform about the attack only afterwards, which is entirely available in the subtitles. The flashy tools and techniques used in password-hacking scenes were more impressive visually, and sometimes the amount of time spent showing this on screen seemed disproportionate to the fraction of subtitles spent describing it. For example, we watched characters bypass the password authentication by using another authentication method (retina scan), brain-to-brain transfer, artificial intelligence as a hacking tool, and fancy illuminated computer screens and tools (without any understandable computer activity shown). However, we found the key points were also understandable without video.

6 Password Strength in Movies

As described in Section 3, we systematically recorded any passwords that were directly stated in scenes ($n = 689$). We analyzed and then evaluated their strength according to *zxcvbn* (Figure 6) and *PGS* (Figure 7). In this section, the results regarding password strength in the movie database will be shown as well as their comparison to real-world passwords.

Per the *zxcvbn* strength metric, well over 70% of movie passwords are rated as Class 0 through Class 2, meaning that they are easily recoverable (i.e., weak). Within these classes, only the 200 Most Popular contains a higher percentage of Class 0 passwords, though this is expected since it by definition excludes uncommon passwords. On the other hand, the two lists that contain real-world passwords, RockYou and Ignis-1M have a much smaller percentage of passwords in Class 0. RockYou has almost twice as many passwords as movie passwords in Class 2 and the same holds for Class 3 passwords. Finally, it is interesting that the two lists that perform best for Class 4 passwords are RockYou and movie passwords. Since RockYou is a complete dataset (the service stored all of its passwords in plaintext) it represents the most accurate distribution of the strength spectrum of real-world passwords and Figure 6 suggests that the percentage of movie passwords classified as Class 4 resembles a real-world distribution very accurately.

Using *PGS*, we see that most passwords fall within the range of 10^3 and 10^8 . The top five strongest passwords from our data set include: `Cv'qrPo` (Our Happy Holiday, 2018), which can not be cracked; `ldfvarumellamsheriaavum`

(Varane Avashyamund, 2020); T19FXP07YT567TZ5 (Those Who Are Fine, 2018); DOOMEDIFYQUQUIT (Sono tornato, 2018); and Youwereeneverthereformed@d (FML, 2016).

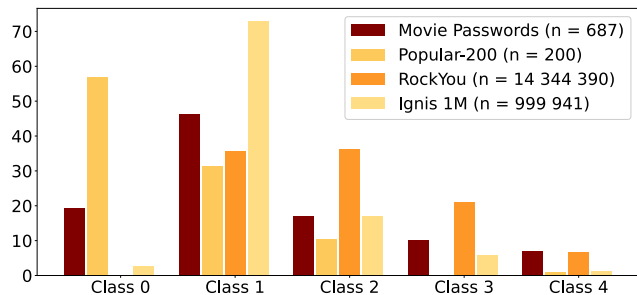


Figure 6: Comparison of the strength distribution of passwords found in movies (movie passwords) and control datasets (Popular-200, RockYou, and Ignis 1M) using the *zxcvbn* classification. The weakest passwords are in Class 0, the strongest in Class 4. The movie passwords contain a distribution closely resembling RockYou with a significant number of passwords belonging in Class 4.

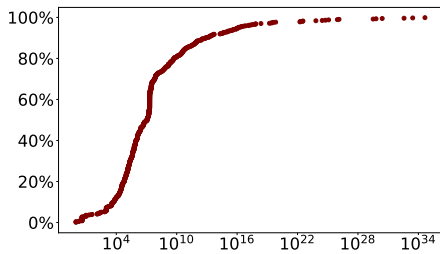


Figure 7: The guessability of passwords found in movies using *PGS*. On the x-axis the number of guesses (log scale) is charted, on the y-axis the percent of how many passwords from within the dataset are guessed.

Are Passwords in Some Contexts Stronger? Returning to our previous context-of-use analysis, we find that the distribution of password strengths is not uniform across all contexts of use (Figure 8). Computer-related and Legitimation passwords have disproportionately more Class 0 passwords, and Internet-related passwords are the context of use with the highest percentage of very strong (Class 4) passwords. Locks, have disproportionately weak passwords – 96% are in Class 0 through 2, which indicates they are easily recoverable, and none are in the strongest Class 4.

There are 73 scenes for which we were able to determine a life cycle point *and* in which a password was directly stated (i.e., for which we can measure the strength of the password in that scene). The majority of passwords in each context had low security (Class 0 or 1, per *zxcvbn*); in password

recovery and reset scenes, 71% were in one of these two lowest-strength classes. The strongest passwords occurred in the context of losing and forgetting a password, where 30% of passwords were Class 3 or 4.

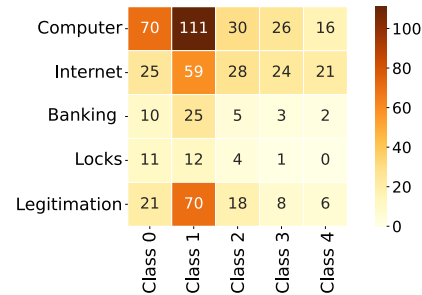


Figure 8: Showcasing the number of cases where password strength according to *zxcvbn* classes overlaps with the different contexts of use. The cell color indicates the relative relationship to other classes by use case with shades of yellow intensifying relative to the code usage in the interval. Strong passwords (Class 3 and 4) are represented more frequently for computer and internet-related topics.

We take a closer look at the more detailed breakdown of Internet-related contexts of use (Figure 9), since they have an especially high percentage of strong passwords. Email and Wi-Fi passwords have an atypical distribution: Passwords of Class 4 are more common than of Class 2 and 3. Streaming/Cable Account passwords stand out as well: The two passwords found for this specific use case are of Class 3 and 4 which makes this category the “strongest” of all categories investigated. However, the small number of passwords found must be taken into account.

This results indicate, that there are certain topics, where strong passwords are considered typical, including email, Wi-Fi, and streaming or cable accounts. In other areas such as locks weak passwords are almost always used.

Are Weaker Passwords More Susceptible to Attack?

133 of the scenes with password guessing (60%) and 10 of the scenes with password hacking (16%) include a specific password whose strength we can analyze. 61 (46%) of password guessing attempts were successful (40% fail and 14% unclear). On the other hand, 100% of the hacking attempts were successful. Both of these success rates are relatively similar to the overall success rates for password guessing and hacking (47% and 81% success, respectively).

Observing how guess success rates differ within the five *zxcvbn* strength classes, as shown in Figure 10, it can be observed that success or failure are only closely related to the strength class. Generally speaking, among low and high classes success or failure of the guessing attack are almost equally likely. Exception is class three where 60% of the attacks are successful.

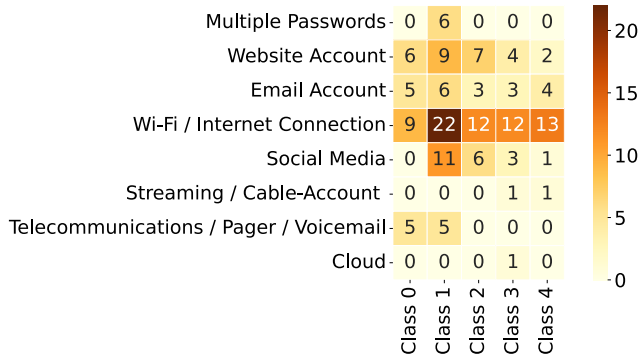


Figure 9: Showcasing the number of cases where password strength according to zxcvbn classes overlaps with the different internet-related contexts of use. The cell color indicates the relative relationship to other classes by use case with shades of yellow intensifying relative to the code usage in the class. Certain usages show a polar distribution of either very strong or very weak passwords.

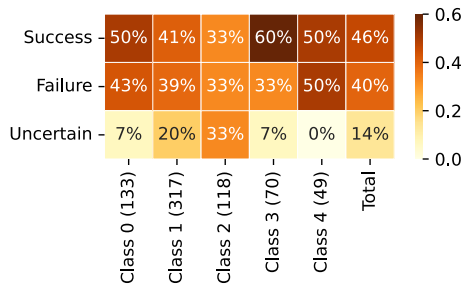


Figure 10: Relation of success, failure and uncertain outcomes of password guessing attacks among different password strength classes according to zxcvbn. The cell color indicates the relative relationship between outcomes of guessing attacks and password class with shades of yellow. It is outstanding that for Class 3, the success of the attack is more likely than its failure.

7 Discussion and Future Work

We found that passwords are increasingly being mentioned in movies (since the start of 2020, around 1 in 20 movies mentioned passwords); thus, it is especially important and timely to consider how realistic they are, what this might be teaching people, and what should be done to create a framework for the dissemination of security topics.

Are Passwords Portrayed Realistically in Movies? Our analysis involved directly comparing the strength of passwords from movies to those leaked in real-world data breaches (see Section 6). Movies contained passwords with a wide range of strengths, just like those in the real-world data sets we compared against. While the patterns we observe depend on exactly which data set we compare to, we found that,

broadly speaking, movies contain more of both especially weak *and* especially strong passwords than most data sets. Interestingly, the fraction of Class 4 passwords (i.e., strongest according to zxcvbn) in movie passwords most closely resembles those in the RockYou data set, which is the only full leaked list of passwords we could compare with, so is in some ways the most realistic.

We can also postulate about the realism of other findings. For example, movies do seem to portray realistic contexts of use (e.g., ranging from car locks or computers to email and streaming accounts), real points in the password life cycle (e.g., creation, change, forgetting, and resetting), and realistic behaviors such as password sharing and reuse. Additionally, the trends in common contexts of use over time seem to have approximately shifted with the evolution of technology: computer-related uses of passwords picked up in the early early 1980s, and Internet-related contexts started catching up or taking over in the mid to late 2010s.

The realism of password hacking and guessing in movies is somewhat of a mixed bag (see Section 5). Out of the total 283 scenes with either type of attack, over half of them were successful which confirms the literature which states that attackers are too powerful [23, 38]. But the *way* password guessing is portrayed is quite realistic: characters use approaches such as trying out typical passwords, using birthdays, hobbies and pet names, or hoping that a person has not changed an old known-to-the-attacker password. Prior work has shown that these types of knowledge about a person can help with guessing their passwords [31]. Guessing often happens in the context of family and friends, which (correctly) shows that attackers are “not only criminal hackers but also people you know” [12]. The small number of attempts needed to guess passwords seems unrealistic, but this may be related to directors’ desire to keep movies from becoming dull.

Unlike prior studies [23], we found that hacking scenes do often have some basis in reality. For example, we found instances of real-world tools such as keyloggers being used in the attack, and we found that attacks exploiting a human factor were common (around 30% of attacks) and more successful than other attack types. Many of the assets targeted in these scenes were also plausible (e.g., company servers as in *Focus* (2015) or illegal platforms investigated by the police as in *The Batman* (2022)). Still, many of the hacking tools and mechanisms shown were unrealistic (e.g., connecting brains, using a hacking artificial intelligence, etc.) and important situations are missing, such as attacking people who are not rich or in particular significant positions, which may lead to the feeling that one is not “important enough to be targeted” [12].

Are Passwords Portrayed Differently in High-Grossing Movies? As stated in Section 4, High-Grossing movies were more likely to contain the word *password*, which is perhaps related to the fact that they are more likely to be newer movies and more likely to be in the genres Thriller, Sci-

Fi and Action & Adventure. Appendix C contains statistics and figures comparing the High-Grossing dataset with the overall dataset. Though we have not performed statistical comparisons, the major patterns are largely consistent with the overall dataset, but there are some interesting differences. There is much more likely to be uncertainty about the success or failure of password guessing or password hacking in High-Grossing movies (around 50%, compared to less than 15% in the overall dataset). Of those scenes where the outcome is known, there is a higher chance in High-Grossing movies that the attack was successful; excluding uncertain outcomes, 80% of hacking and guessing attempts in High-Grossing movies are successful compared to only 62% in the full dataset. High-grossing movies are much less likely to depict the Change or Training phases of the password Life Cycle (18% compared to 44% in the overall dataset), much more likely to show a password Reset (24% compared to 8% in the overall dataset), and somewhat more likely to show password Creation (29% compared to 21% in the overall dataset). None of the seven movies that contained multi-factor authentication were in this High-Grossing dataset, nor was the one movie that showed a password manager.

What Are People Likely to Be Learning from Movies?

While our study scope is focused on the contents of movies, we know from prior work that movies can influence viewers' understanding of cybersecurity topics [4, 20, 73]. Our findings suggest that there are both good and bad security and privacy practices in movies that viewers may be learning from, and we find that some security best practices are rarely shown (i.e., it is implausible that viewers would learn to follow these based on watching movies).

We focus here on two key positive practices that viewers could take away from the movies in our data set. First, even though many of the passwords included in movies are somewhat weak, contain personal details such as birth dates that are known to make passwords more guessable [31] and/or are easily guessed by characters, we have hope that many of these scenes may actually be teaching viewers about the characteristics of weak passwords. Most simply, when a weak password is shown as being easily guessed, perhaps this is a cue to viewers that the password is weak. We also observed that characters often make fun of bad passwords in movies, which provides even more direct commentary to viewers. Second, there *are also* many strong passwords in movies. We found that the strongest passwords were used in Internet-related contexts. Within this relatively broad category, strong passwords were especially commonly used for email, Wi-Fi, and streaming or cable accounts. We expect that these scenes could help normalize the use of strong passwords.

As expected, many aspects of our findings point to troublesome lessons viewers might glean from movies. Password sharing and reuse are portrayed as normal behaviors, even super villains use 12345 as their password, and even the boss

of an IT company puts his password under the keyboard. The inclusion of some of these in movies may be justifiable – for example, password sharing involves interaction between people, which plays well to getting multiple characters involved in a scene, but movies rarely spend much time following one character all alone and doing things that might involve passwords (which we must admit, are quite boring). However, even if these behaviors *are* realistic, normal and somewhat justifiable, it is likely still harmful for viewers to see them normalized in movies. Some of the weakest passwords were used for locks or banking, which are high-risk contexts (in fairness, many of these were number PINs, which are realistic and more guessable due to their short length).

Returning to the overall unrealistically high success rate of hacking or guessing passwords in movies, which is even higher in High-Grossing movies, we found that the strength of a password has practically no effect on security. Strong passwords are guessed just as often as weak ones (or even more often) and even for the most secure passwords of Class 4 that are very difficult to guess in the real world, passwords such as *Stephanie'sdude2016* are guessed correctly within seconds. Combined, these portrayals might send the message to viewers that attacks will be successful regardless of security efforts, so why bother trying? In High-Grossing movies, we observed that the outcome of a hacking or guessing attempt was more often uncertain compared to the baseline. We hypothesize that this could influence viewers to see cybersecurity as unapproachable and mysterious, further contributing to tendencies to avoid learning about it and taking appropriate security measures in their real lives.

Finally, we found fewer than 0.3% of movies that mention the use of password managers and/or multi-factor authentication, which are widespread and commonly suggested as part of security best practices [2]. This presents a missed opportunity for movies to help familiarize viewers with these tools.

Implications for Film and Policy Makers, Educators, and Researchers

Because our findings show that many movies portray passwords in ways that could lead viewers to riskier security and privacy behaviors, this paper underscores the importance of recommendations from prior work that call for the creation of a “Cybersecurity in Entertainment Task Force” to consult with both security experts and film makers to help ensure that portrayals of passwords (or of technology more broadly) does not lead to harmful negative outcomes for viewers [20]. Such consulting efforts have already been successful in other domains, such as medicine. Additionally, we contribute an understanding of what contexts of use, password-related behaviors, plot dynamics, and misleading or problematic portrayals have been most common in movies so far. This could help consultants tailor what topics they are most prepared to consult on, and it could help them guide film makers to decisions that are less cliché.

Prior work has also emphasized the capacity for educators to leverage movie scenes in their lessons [20]. We agree and suggest that these could help engage students and enhance their understanding of the content. Studies have shown that the inclusion of movie clips in other educational is promising [18, 34, 39], but should be approached with care when the clips contain inaccuracies [10]. As stated in Section 3, we released a database of our findings (i.e., codes) for each movie in our data set. This can be a helpful resource for cybersecurity educators to find the most relevant clips. Additionally, our findings can help guide curriculum development and instructor decisions regarding which topics are most important to cover (i.e., perhaps focusing on topics that our work suggests students are especially likely to be misinformed about).

Finally, our findings motivate future work in this research direction. As discussed in our limitations section, it is likely worthwhile to study other forms of entertainment media in similar ways. For example, television shows (or series) follow the same characters over longer periods and, we imagine, are more likely to include scenes with normal, everyday uses of passwords. The relative normalcy of life on television (compared to in movies) might make these portrayals seem more realistic to viewers. Our work did not provide insights on *how* to expand the scope of analysis beyond the topic of passwords; however, our findings demonstrate that doing so could help solidify our knowledge about how cybersecurity and privacy topics (or technology more broadly) are portrayed in movies and, thus, what viewers might be learning. Along these lines, it was beyond the scope of our study to determine how the elements of movies we identified actually impact people, but this is an important next step. Finally, while we have suggested above that instructors could incorporate movie clips in their classes, other fields where this is common have conducted studies to understand how this should best be done and how to avoid common pitfalls; this type of follow-on work would be beneficial in this domain as well.

8 Conclusion

To analyze the depiction of passwords and password behavior in movies, we performed a subtitle analysis and watched selected scenes. Our results show a broad spectrum of different password activities and contexts of usage in movies from various years and genres. Movies show passwords of different strengths and outline different kinds of password attacks. However, the chances of success are presented as dangerously high and important best practices are missing from the portrayal. We aim to contribute towards a better understanding of how cybersecurity is depicted in the media, and ultimately to a better understanding of how we can mitigate the (negative) consequences of wrongful depiction of cybersecurity.

Acknowledgments

We want to acknowledge and thank the many people who contributed to this work over a long period of time. Tadayoshi Kohno and Alexis Hiniker provided ideas and feedback in the very early stages. Clemend Zhong, Saloni Vaishnav, and Effie Karas did REU projects related to this work that informed the final study design. Maximilian Golla contributed vital support with the data set. Tobias Hägele, Stina Schäfer, Sarina Javdani and Daniel Janßen contributed to data analysis, including viewing movies, and data visualization. Andrea Watkins contributed feedback on an earlier draft. We thank the anonymous shepherd for the invaluable help in getting the paper ready to be published.

References

- [1] Ruba Abu-Salma, M. Angela Sasse, Joseph Bonneau, Anastasia Danilova, Alena Naiakshina, and Matthew Smith. Obstacles to the Adoption of Secure Communication Tools. In *IEEE Symposium on Security and Privacy, SP '17*, pages 137–153, San Jose, California, USA, May 2017. IEEE.
- [2] Yusuf Albayram, John Liu, and Stivi Cangonj. Comparing the Effectiveness of Text-based and Video-based Delivery in Motivating Users to Adopt a Password Manager. In *European Workshop on Usable Security, EuroUSEC '21*, pages 89–104, Virtual Conference, October 2021. ACM.
- [3] Anonymous Reddit User. r/DataHoarder: 5,719,123 Subtitles From OpenSubtitles.org, July 2022. https://www.reddit.com/r/DataHoarder/comments/w7sgcz/5719123_subtitles_from_opensubtitlesorg/, as of June 6, 2024.
- [4] Khadija Baig, Elisa Kazan, Kalpana Hundlani, Sana Maqsood, and Sonia Chiasson. Replication: Effects of Media on the Mental Models of Technical Users. In *Symposium on Usable Privacy and Security, SOUPS '21*, pages 119–138, Virtual Conference, August 2021. USENIX.
- [5] Travis Bell and Community. The Movie Database: Popularity & Trending, November 2023. <https://developer.themoviedb.org/docs/popularity-and-trending>, as of June 6, 2024.
- [6] Pablo González Blasco. Literature and Movies for Medical Students. *Family Medicine*, 33(6):426–428, June 2001.
- [7] BonaFideBoss. IMDb: Top 1000 Highest-Grossing Movies of All Time, April 2024. <https://www.imdb.com/list/ls098063263/>, as of June 6, 2024.

- [8] “Bran0” and Community. Open Subtitles: Download Movie and TV Series Subtitles, January 2006. <https://www.opensubtitles.org>, as of June 6, 2024.
- [9] Ulla Bunz. “We speak in code, in case the telephone operator should be eavesdropping!”: How Popular Movies Reflect Society’s Attitude Toward Technology. In *Annual Convention of the Media Ecology Association*, MEA ’03, pages 1–18, Hempstead, New York, USA, June 2003. MEA.
- [10] Andrew C. Butler, Franklin M. Zaromb, Keith B. Lyle, and Henry L. Roediger. Using Popular Films to Enhance Classroom Learning: The Good, the Bad, and the Interesting. *Psychological Science*, 20(9):1161–1168, September 2009.
- [11] Sonia Chiasson and Paul C. Van Oorschot. Quantifying the Security Advantage of Password Expiration Policies. *Designs, Codes and Cryptography*, 77(2–3):401–408, December 2015.
- [12] Mathieu Christmann, Peter Mayer, and Melanie Volkamer. Vision: What Johnny learns about Password Security from Videos posted on YouTube. In *European Workshop on Usable Security*, EuroUSEC ’21, pages 124–128, Virtual Conference, October 2021. ACM.
- [13] Wikipedia Community. Wikipedia: Film analysis, May 2024. https://en.wikipedia.org/w/index.php?title=Film_analysis&oldid=1213215167, as of June 6, 2024.
- [14] Nik Cubrilovic. RockYou Hack: From Bad To Worse, December 2009. <https://techcrunch.com/2009/12/14/rockyou-hack-security-myspace-facebook-passwords/>, as of June 6, 2024.
- [15] Maria Dimitropoulou, Jon Andoni Duñabeitia, Alberto Avilés, José Corral, and Manuel Carreiras. Subtitle-Based Word Frequencies as the Best Estimate of Reading Behavior: The Case of Greek. *Frontiers in Psychology*, 1:218:1–218:12, December 2010.
- [16] “Fast Eddie” and Community. TV Tropes Pop-Culture Wiki: Hollywood Hacking, November 2010. <https://tvtropes.org/pmwiki/pmwiki.php/Main/HollywoodHacking>, as of June 6, 2024.
- [17] “Fast Eddie” and Community. TV Tropes Pop-Culture Wiki: Hollywood Encryption, January 2014. <https://tvtropes.org/pmwiki/pmwiki.php/Main/HollywoodEncryption>, as of June 6, 2024.
- [18] Stefan Engert and Alexander Spencer. International Relations at the Movies: Teaching and Learning about International Politics through Film. *Perspectives: Review of International Affairs*, 17(1):83–103, July 2009.
- [19] Franklin Fearing. Influence of the Movies on Attitudes and Behavior. *The Annals of the American Academy of Political and Social Science*, 254:70–79, November 1947.
- [20] Kelsey R. Fulton, Rebecca Gelles, Alexandra McKay, Yasmin Abdi, Richard Roberts, and Michelle L. Mazurek. The Effect of Entertainment Media on Mental Models of Computer Security. In *Symposium on Usable Privacy and Security*, SOUPS ’19, pages 79–95, Santa Clara, California, USA, August 2019. USENIX.
- [21] Brian Gallagher. 10 Great Social Commentary Movies That Reflect Contemporary Society, August 2017. <https://www.tasteofcinema.com/2017/10-great-social-commentary-movies-that-reflect-contemporary-society/>, as of June 6, 2024.
- [22] Maximilian Golla and Markus Dürmuth. On the Accuracy of Password Strength Meters. In *ACM Conference on Computer and Communications Security*, CCS ’18, pages 1567–1582, Toronto, Ontario, Canada, October 2018. ACM.
- [23] Damian Gordon. Forty Years of Movie Hacking: Considering the Potential Implications of the Popular Media Representation of Computer Hackers from 1968 to 2008. *International Journal of Internet Technology and Secured Transactions*, 2(1/2):59–87, February 2010.
- [24] Hana Habib, Pardis Emami Naeini, Summer Devlin, Maggie Oates, Chelse Swoopes, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. User Behaviors and Attitudes Under Password Expiration Policies. In *Symposium on Usable Privacy and Security*, SOUPS ’18, pages 13–30, Baltimore, Maryland, USA, August 2018. USENIX.
- [25] Ata Hakçıl (“ignis sec”). PWDB: New Generation of Password Mass-Analysis, July 2020. <https://github.com/ignis-sec/Pwdb-Public>, as of June 6, 2024.
- [26] Cormac Herley and Paul C. Van Oorschot. A Research Agenda Acknowledging the Persistence of Passwords. *IEEE Security & Privacy*, 10(1):28–36, January 2012.
- [27] Beth L. Hoffman, Ariel Shensa, Charles Wessel, Robert Hoffman, and Brian A. Primack. Exposure to Fictional Medical Television and Health: A Systematic Review. *Health Education Research*, 32(2):107–123, April 2017.
- [28] Lee Humphreys, Phillipa Gill, and Balachander Krishnamurthy. Twitter: A Content Analysis of Personal Information. *Information, Communication & Society*, 17(7):843–857, October 2013.

- [29] Friska Sari Luksiana Hutajulu and Herman Herman. Analysis of Illocutionary Act in the Movie “You Are My Home” English Subtitle. *Journal of English Educational Study*, 2(1):29–36, May 2019.
- [30] Philip Nicholas Johnson-Laird, Vittorio Girotto, and Paolo Legrenzi. *Mental Models: A Gentle Guide for Outsiders*, April 1998. <https://web.archive.org/web/20050305184203/https://www.si.umich.edu/ICOS/gentleintro.html>, as of June 6, 2024.
- [31] Aikaterini Kanta, Iwen Coisel, and Mark Scanlon. A Novel Dictionary Generation Methodology for Contextual-Based Password Cracking. *IEEE Access*, 10:59178–59188, June 2022.
- [32] Aikaterini Kanta, Iwen Coisel, and Mark Scanlon. Harder, Better, Faster, Stronger: Optimising the Performance of Context-Based Password Cracking Dictionaries. *Forensic Science International: Digital Investigation*, 44:301507:1–301507:9, March 2023.
- [33] Aikaterini Kanta, Sein Coray, Iwen Coisel, and Mark Scanlon. How Viable Is Password Cracking in Digital Forensic Investigation? Analyzing the Guessability of over 3.9 Billion Real-World Accounts. *Forensic Science International: Digital Investigation*, 37:301186:1–301186:11, July 2021.
- [34] Gary Koch and Colette T. Dollarhide. Using a Popular Film in Counselor Education: Good Will Hunting as a Teaching Tool. *Counselor Education and Supervision*, 39(3):203–210, March 2000.
- [35] Herbert E. Krugman and Eugene L. Hartley. Passive Learning From Television. *The Public Opinion Quarterly*, 34(2):184–190, June 1970.
- [36] Marcus Law, Wilson Kwong, Farah Friesen, Paula Veinot, and Stella L. Ng. The Current Landscape of Television and Movies in Medical Education. *Perspectives on Medical Education*, 4(5):218–224, September 2015.
- [37] Pierre Lison and Jörg Tiedemann. Opensubtitles2016: Extracting large parallel corpora from movie and tv subtitles. *10th International Conference on Language Resources and Evaluation (LREC 2016)*, pages 923–929, 2016. <https://opus.nlpl.eu/OpenSubtitles/corpus/version/OpenSubtitles>, as of June 6, 2024.
- [38] Johnny Long (“j0hnnny”). DEFCON 14: Secrets of the Hollywood Hacker!, August 2006. https://www.youtube.com/watch?v=m_Xmc49ZrYA, as of June 6, 2024.
- [39] Abolfaz Mahdiloo and Siros Izadpanah. The Impact of Humorous Movie Clips on Better Learning of English Language Vocabulary. *International Journal of Research in English Education*, 2(2):16–30, June 2017.
- [40] Paweł Mandra, Emmanuel Keuleers, Zofia Wodniecka, and Marc Brysbaert. Subtlex-Pl: Subtitle-Based Word Frequency Estimates for Polish. *Behavior Research Methods*, 47(2):471–483, June 2015.
- [41] Peter Mayer, Collins W. Munyendo, Michelle L. Mazurek, and Adam J. Aviv. Why Users (Don’t) Use Password Managers at a Large Educational Institution. In *USENIX Security Symposium, SSYM ’22*, pages 1849–1866, Boston, Massachusetts, USA, August 2022. USENIX.
- [42] Peter Mayer and Melanie Volkamer. Addressing Misconceptions about Password Security Effectively. In *Workshop on Socio-Technical Aspects in Security and Trust, STAST ’17*, pages 16–27, Orlando, Florida, USA, December 2017. ACM.
- [43] Daniel Miessler and Community. SecLists: “200 Most Used Passwords”, December 2023. https://github.com/danielmiessler/SecLists/blob/master/Passwords/2023-200_most_used_passwords.txt, as of June 6, 2024.
- [44] Hani Morgan. Counteracting Misconceptions about the Arab World from the Popular Media with Culturally-Authentic Teaching. *International Social Studies*, 2(2):70–83, January 2013.
- [45] National Cyber Security Centre. The Problems with Forcing Regular Password Expiry, December 2016. <https://www.ncsc.gov.uk/articles/problems-forcing-regular-password-expiry>, as of June 6, 2024.
- [46] Katharina Pfeffer, Alexandra Mai, Edgar Weippl, Emilee Rader, and Katharina Krombholz. Replication: Stories as Informal Lessons about Security. In *Symposium on Usable Privacy and Security, SOUPS ’22*, pages 1–18, Boston, Massachusetts, USA, August 2022. USENIX.
- [47] McKenna Prancing. I Was a Medical Advisor for Grey’s Anatomy. Here’s What I Learned., October 2017. <https://rightasrain.uwmedicine.org/well/stories/i-was-medical-advisor-greys-anatomy-heres-what-i-learned>, as of June 6, 2024.

- [48] Emilee Rader, Rick Wash, and Brandon Brooks. Stories as Informal Lessons about Security. In *Symposium on Usable Privacy and Security*, SOUPS '12, pages 6:1–6:17, Washington, District of Columbia, USA, July 2012. ACM.
- [49] Pritham Y. Raj. Medicine, Myths, and the Movies. Hollywood's Misleading Depictions Affect Physicians, Patients Alike. *Postgraduate Medicine*, 113(6):9–10, June 2003.
- [50] Mudassar Raza, Muhammad Iqbal, Muhammad Sharif, and Waqas Haider. A survey of password attacks and comparative analysis on methods for secure authentication. *World applied sciences journal*, 19(4):439–444, 2012.
- [51] Elissa M. Redmiles, Sean Kross, and Michelle L. Mazurek. How I Learned to Be Secure: A Census-Representative Survey of Security Advice Sources and Behavior. In *ACM Conference on Computer and Communications Security*, CCS '16, pages 666–677, Vienna, Austria, October 2016. ACM.
- [52] Elissa M. Redmiles, Amelia R. Malone, and Michelle L. Mazurek. I Think They're Trying to Tell Me Something: Advice Sources and Selection for Digital Security. In *IEEE Symposium on Security and Privacy*, SP '16, pages 272–288, Los Alamitos, CA, USA, May 2016. IEEE Computer Society.
- [53] "Rhiannon". Hacker's Game: 10 Things Hollywood Got Wrong About Computer Hacking, July 2022. <https://hotbotvpn.com/blog/10-things-hollywood-got-wrong-about-computer-hacking/>, as of June 6, 2024.
- [54] Scott Ruoti, Tyler Monson, Justin Wu, Daniel Zappala, and Kent Seamons. Weighing Context and Trade-offs: How Suburban Adults Selected Their Online Security Posture. In *Symposium on Usable Privacy and Security*, SOUPS '17, pages 211–228, Santa Clara, California, USA, July 2017. USENIX.
- [55] Richard Shay, Saranga Komanduri, Adam L. Durity, Phillip (Seyoung) Huh, Michelle L. Mazurek, Sean M. Segreti, Blase Ur, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. Can long passwords be secure and usable? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 2927–2936, New York, NY, USA, April 2014. Association for Computing Machinery.
- [56] University of North Carolina at Chapel Hill The Writing Center. Film analysis, May 2024. <https://writingcenter.unc.edu/tips-and-tools/film-analysis/>, as of June 6, 2024.
- [57] Ghent University. DICIS – A Scientific Research Network on Digital Cinema Studies, May 2024. <https://www.ugent.be/ps/communicatiewetenschappen/cims/en/research/current-research-projects/dicis.htm>, as of June 6, 2024.
- [58] Blase Ur, Jonathan Bees, Sean M. Segreti, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. Do Users' Perceptions of Password Security Match Reality? In *ACM Conference on Human Factors in Computing Systems*, CHI '16, pages 3748–3760, San Jose, California, USA, May 2016. ACM.
- [59] Blase Ur, Sean M. Segreti, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, Saranga Komanduri, Darya Kurilova, Michelle L. Mazurek, William Melicher, and Richard Shay. Measuring Real-World Accuracies and Biases in Modeling Password Guessability. In *USENIX Security Symposium*, SSYM '15, pages 463–481, Washington, District of Columbia, USA, August 2015. USENIX.
- [60] USC Annenberg Norman Lear Center. Hollywood, Health and Society: About Us, February 2024. <https://hollywoodhealthandsociety.org/about-us/>, as of June 6, 2024.
- [61] Niklas von Boguszewski, Sana Moin, Anirban Bhowmick, Seid Muhie Yimam, and Chris Biemann. How Hateful are Movies? A Study and Prediction on Movie Subtitles. *CoRR*, abs/2108.10724:1–12, August 2021.
- [62] Ding Wang, Ping Wang, Debiao He, and Yuan Tian. Birthday, Name and Bifacial-Security: Understanding Passwords of Chinese Web Users. In *USENIX Security Symposium*, SSYM '19, pages 1537–1555, Santa Clara, California, USA, August 2019. USENIX.
- [63] Rick Wash. Folk Models of Home Computer Security. In *Symposium on Usable Privacy and Security*, SOUPS '10, pages 11:1–11:16, Redmond, Washington, USA, July 2010. ACM.
- [64] OpenSubtitles Webmasters. Disclaimer - opensubtitles.org, May 2024. <https://www.opensubtitles.org/de/disclaimer>, as of June 6, 2024.
- [65] Danny Wedding and Ryan M. Niemiec. *Movies and Mental Illness: Using Films to Understand Psychopathology*. Hogrefe Publishing, Göttingen, Germany, 3 edition, 2009.
- [66] Daniel Lowe Wheeler. zxcvbn: Low-Budget Password Strength Estimation. In *USENIX Security Symposium*, SSYM '16, pages 157–173, Austin, Texas, USA, August 2016. USENIX.

- [67] Naomi White. How Hollywood Movies Reflect Society, September 2022. <https://www.thegreatdebatersmovie.com/how-hollywood-movies-reflect-society/>, as of June 6, 2024.
- [68] David Knapp Whittier, May G. Kennedy, Janet S. St. Lawrence, Salvatore Seeley, and Vicki Beck. Embedding Health Messages into Entertainment Television: Effect on Gay Men’s Response to a Syphilis Outbreak. *Journal of Health Communication*, 10(3):251–259, September 2005.
- [69] Tannis Macbeth Williams. How and What Do Children Learn from Television? *Human Communication Research*, 7(2):180–192, December 1981.
- [70] Muheng Yu, Michael C. Carter, Drew P. Cingel, and Jeanette B. Ruiz. A Content Analysis of Aggression in Netflix Original, Adolescent-Directed Series’ Subtitles. *Communication Quarterly*, 71(5):588–609, August 2023.
- [71] Muheng Yu, Michael C. Carter, Drew P. Cingel, and Jeanette B. Ruiz. How Sex Is Referenced in Netflix Original, Adolescent-Directed Series: A Content Analysis of Subtitles. *Psychology of Popular Media*, 13(1):1–11, January 2024.
- [72] Samira Zibaei, Dinah Rinoa Malapaya, Benjamin Mercier, Amirali Salehi-Abari, and Julie Thorpe. Do Password Managers Nudge Secure (Random) Passwords? In *Symposium on Usable Privacy and Security*, SOUPS ’22, pages 581–597, Boston, MA, USA, August 2022. USENIX.
- [73] Verena Zimmermann and Nina Gerber. “If It Wasn’t Secure, They Would Not Use It in the Movies” - Security Perceptions and User Acceptance of Authentication Technologies. In *Human Aspects of Information Security, Privacy and Trust*, HAS ’17, pages 265–283, Vancouver, British Columbia, Canada, July 2017. Springer.

A Passwords in Movies: List and Strength Analysis Results with *zxcvbn*

Passwords were repeatedly seen in the analyzed movies. With *zxcvbn* these are sorted into five Strength Categories. Below we show excerpts from the list. The complete list can be viewed online³.

³<https://www.itsec.uni-hannover.de/de/usec/forschung/medien/password-depiction-in-movies>

Class 0 (133) 0000, 01234, 0515, 1111, 1212, 123, 123123, 123321, 1234, 12345, 123456, 164, 179, 1951, 1967, 1972, 1982, 1998, 2222, 2345, 2468, 286, 314, 314159, 326, 38, 4040, 4321, 437, 438, 500, 521, 651, 680, 69, 696969, 761, 77777, 923, 949, 999999, 999999999, a2h, ABC123, ABCD1234, Angela, Anna, Annie, Barbara, Batman, beer, Birdie, Bob, Boobs, butterfly, carmen, Casper, Crystal, Denise, diamond, Die, eat, Enter, Eric, Erica, erin, Faye, Frankie, freedom, girls, guess, guest, h0us3, Heaven, Horny, James, Jenny, Justin, Laura, leon, Love, Lucas, March, Melanie, Mountain, Myself, Natasha, nose, Om, Orlando, Paradise, Party, password, Password, PASSWORD, Peaches, Pedro, Pepper, pirate, Porn, princess

Class 1 (317) 0113, 040515, 0511, 0512, 05171210, 0522, 0623, 0627, 070476, 0708, 0710, 072099, 0801, 1048, 1104, 1112, 1126, 1166, 1192, 1195, 1230, 1321, 132109, 1356, 1492, 15626, 1685198, 1776, 1796, 19300830, 19891023, 20107, 20131026, 2111, 22093, 2235, 2259, 2356, 2372, 2501, 2598, 262670, 27130, 295141, 2QUILA, 3041, 3057, 3690, 4093, 420God, 4664, 489*48, 4989, 5023, 5042, 5321, 541267, 5445, 54AGT, Survivor, 6143, 6246, 627628, 661968, 691234, 712735, 7232, 7397, 7590, 8224, 8644, 8854, 977127, 9993, a/321, A3501, Abby, Abdel, adventure, AirBud, Amen, Angelo, Angelangel, Angiovanni, annie123, Anusua, argonaut, athlete, Autumnleave, Bacon, badmama, baloney1, Barnsey, Bassola, Bastard, beagles, BEARD, Beethoven, Begood, Belle1998, BigBen, Biggie, BlackChicken, BlackOut, bluebeauty, Bondik, Boomerang, boxing, BRANGELINA, Brat, Briefs, Brigitte, Buremma, buttercup, Cancerian, Carmim, casket, catnip, catnip1, chandelier, Charmer, Chestnut, chicken65, Chicken65, ChiefAsshole, ChowChow, chrisnewton, CM110

Class 2 (118) 0505informer, 1/2-1/2, 14-J-89, 2060Pinto, 2516904, 801023000000, ACAPULCO01, asavari, asstastic, ATR1020, auroraborealis, AVCHomes, Ayla123, B055man69, badmamma, Balki1987, BaluMama, Bankerchick, Batfan1, bayernmunich, Beatrix928, BettyGrable, BODYGARD, canttell, Carlton071133, CarryGold, Césoul89, Charbear, Chewinggum, Clavius, Cloudberry, Creamcrackers, cuddlefresh, Cv’qrPo, Damnedmelon, darlinggoli, ddayspm, DEADRIPLEY, DeathWhisper, DevAnand, DJDESFAS, Doorlogs, ExtraStuffing, Fartnoise, gindrick, Haircomb, HAPPYMANPAN, HDA14+1, Helvetica, hoodfume, HumphreyBogart, IAN&EMMA, ilikelaura, imthman, interzone

Class 3 (70) 13C34RMXL, 2015salesstrategy, 45gx67kn21, 4saraandjimmy, 68k305RW65, Abhimanyu, Abraxas79713, arthurisadick, asami0709, Barbsguy1989, bardahlia13, batchoy0706, bauer-smythe, bethmarch4eva, Blumenfeld, boobfart69, broccoli34525A, ch3ryjone3s, cocknballs, daddysprincess1994, DavidFosterWallace, DirtyDaniels69, dividebyzero, Dongmaster82, Donkeyballs84, Effenberg, Eightclap1, Elchapo69, fluxcapacitor, GOD’S GIFT, Grid90245, Grilledcarrots, Hananamiti, Hasselhof, Heaven’sDoor, icantellyou, ILoveYudi, Indiansubcontinent79, Johnnyutah69, Konigshutte, liz0919/85, louisepaul222, MaRc62?!*\$, meganthe moron, milkandcookie\$, MillerEmployeeGeneral, minayo0118, misterdarcyforever

Class 4 (49) 72435637440472, Aatukaalamma, alfreddabuttler, AndrétheGiant, arianagrandespuffynipples555, Asagolabius, Bagofdicks44, Baitursinov, Bergen-Belsen, BickmanGuest, bigbertandsmallbert, bryansbabydick69, CafeBonaparte, Cantarpiano1863, catwomanisaBitch, DabanggSultan, DOOMEDIFYQUQUIT, Epluribusfunk, EvianBottledAir, GabriMarta202, HAWTHRONE1850, Heymonaumona, ILOVEchotaBHEEM143, itsagratefuldead, JoyMukherji, k!TTeN!ckler312, Kavya_Kavya, ldfvarumellamsheriaavum, Marsupilami, MeikoMochizuki, MilkyShonku21, ninelformiguel75, OffWithTheirHeads!, penis_grigio72, piazzadellecinquelune, Prabhavathi, ROSEPOGONIAS, StarBigSkyChristmas!, Stephanie’sdude2016

B Points in the Password Life Cycle: Descriptions and Sample Scenes

During the password life cycle different activities are performed, as described in Table 1.

Life cycle point	Description	Example
Password Creation	Setting up a password or speaking about password generation	“He said that he would create a website. In order to access the website, I would need a password. So he took a paper napkin that was on the table in this cafe where we were talking in Brussels and he hooked together several of the words in the commercial logo [...]” (We Steal Secrets: The Story of WikiLeaks, 2013)
Password Change	Changing the password as account owner or legitimated person or intending to	“To flush out the mole is easy, change our password and signals, tell all the others, and pretend nothing’s wrong.” (The Swordswoman in White, 1992)
Training to Remember a Password	Checking if someone else still remembers the password or reminding them of it	“I’ll see you in an hour? -Right. -Haven’t forgotten the password? -Whatever gave you the idea?” (The Body, 2003)
Password Loss and Oblivion	Failing to remember the password or losing a physical reminder like a piece of paper	“My favorite is when they come in, forgotten their password, Locked themselves out of their own computer.” (The Zombie Werewolves Attack!, 2009)
Password Recovery, Reset, and Hints	Changing the password using recovery systems or receiving hints to remember the password	“All that you do is enter an email address and attempt to enter a password. Then, you see, it asks if you forgot your password. So you click that and it tells you to check your email to change your password. So then I go to her email [...] (16 and Missing, 2015)
Password Reuse	Reusing the same password for multiple different purposes or accounts	“will need the passwords to your email accounts, your social media accounts, your bank accounts, your credit card accounts and your Cinnabon Rewards account. - It’s easy. It’s the same password for all of ’em. It’s phil123456. - You’ve got to be kidding me.” (Jexi, 2019)

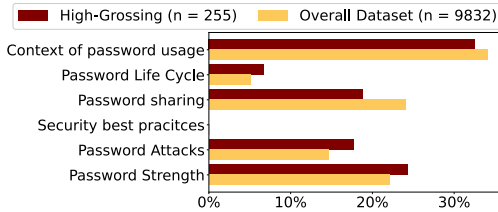
Table 1: Password activities within the movies: Activity names, descriptions, and example scenes.

C Comparison of the Dataset with High-Grossing Movies

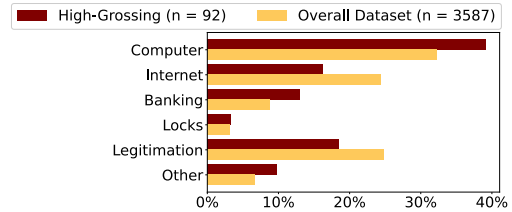
High-grossing movies containing the word “password” (70): In the *Top 1000 Highest-Grossing Movies of all Time* [7] we identified 70 movies containing the word *password* at least once (see Section 3). Those are:

Alice in Wonderland; Armageddon; Avengers: Endgame; Batman Returns; Bruce Almighty; Captain America: The Winter Soldier; Captain Marvel; Captain Phillips; Casino Royale; Cheaper by the Dozen; Crazy Alien; Crazy Rich Asians; Disclosure; Doctor Strange; Elysium; Ghost; Godzilla vs. Kong; GoldenEye; Harry Potter and the Chamber of Secrets; Harry Potter and the Prisoner of Azkaban; Harry Potter and the Sorcerer’s Stone; Heat; Home; Ice Age: Continental Drift; Iron Man 3; It; It Chapter Two; Kingsman: The Secret Service; Lucy; Men in Black: III; Mojin: The Lost Legend; Monster Hunt; National Treasure; Ne Zha; Non-Stop; Now You See Me; Parasite; Pitch Perfect 2; Ralph Breaks the Internet; Ready Player One; Safe House; Sex and the City; Spider-Man: Far from Home; Spider-Man: Into the Spider-Verse; Superman Returns; Tangled; Terminator Genisys; The Batman; The Bodyguard; The Break-Up; The Departed; The Emoji Movie; The Firm; The Hangover Part II; The Hangover Part III; The Hitman’s Bodyguard; The Incredibles; The Intern; The Lego Batman Movie; The Lord of the Rings: The Fellowship of the Ring; The Other Woman; The Pacifier; The Secret Life of Pets; The Shape of Water; The Social Network; The SpongeBob Movie: Sponge Out of Water; The Vow; Tomorrowland; True Lies; Who Framed Roger Rabbit

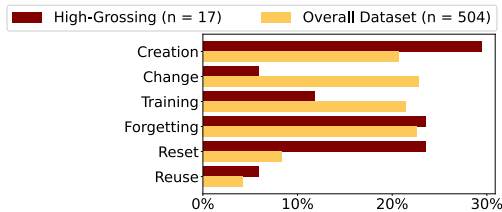
Comparative Statistics: For each topic of the paper, the code frequencies were calculated for the high-grossing movies and the entire data set. The distribution within the different topics is compared in Figure 11 and in Figure 12 the differences in attacks are shown.



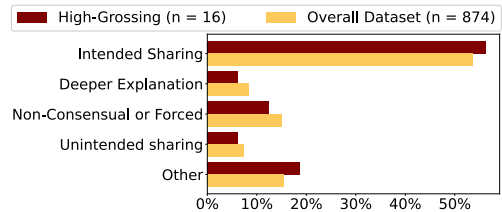
(a) **Overview of topic distribution** This indicator of which topics occur rather frequently or rarely within the data set shows, that basically, all topics except best practices are present in both data sets and also with approximately similar distribution.



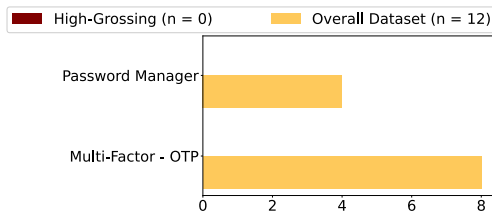
(b) **Context of Use** As described in Section 4.2, it is in often identifiable what the password is used for. In the high-grossing movies there are more movies used in the compute context and less regarding Internet and Legitimation.



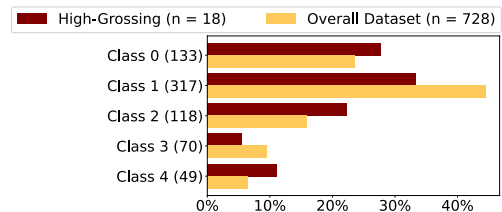
(c) **Life Cycle** The different scenarios in the life cycle of passwords (cf. Section 4.2) occur in the High-Grossing Movies but with different characteristics. Only one movie contains Change and Reuse, only two Training.



(d) **Password sharing** In 16 scenes in the top-grossing movies password sharing is presented. The frequency distribution is very similar to that of the full dataset. Similar as in the whole dataset, most common type is *Intended/Intentional* sharing.

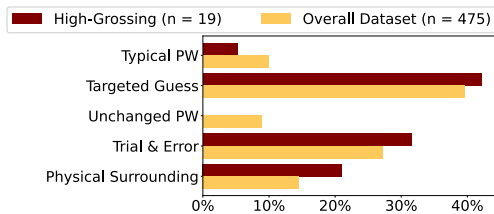


(e) **Security Best Practices in total numbers** This topic appears only in the overall dataset and not in the top-grossing movies. The number of scenes (12) is negligible (0.12% of codes in the set of all codes applied to the overall dataset).

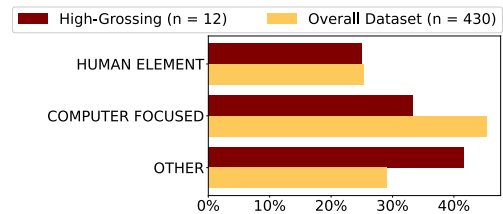


(f) **Password Strength: zxcvbn** Passwords with all strength categories using zxcvbn appear in both the high-grossing and the overall dataset. High-grossing movies contain fewer class 1 passwords than the overall dataset.

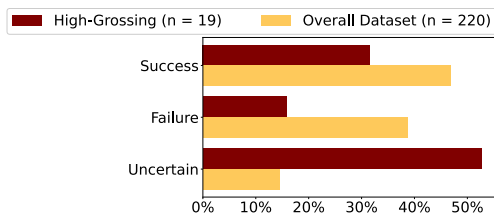
Figure 11: Comparison of high-grossing and overall dataset for different topics presented in this paper. The number of codes per topic was calculated for each diagram for both the high-grossing and the overall dataset and was used to quote the distribution.



(a) **Password Guessing: Approach** In the top-grossing movies, unlike in the overall dataset, no unchanged password is used for guessing. Otherwise, the frequency of the procedures is similar.

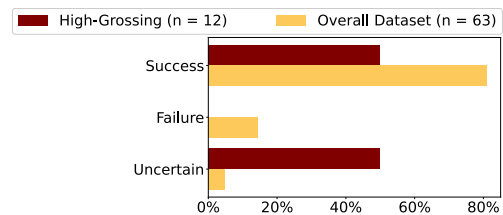


(b) **Password Hacking: Approach** There are only 12 scenes, therefore rough categories are used. The high-grossing movies contain more "other" and less computer-based approaches.



(c) **Password Guessing: Success** For the high-grossing movies, an above-average number of the guessing approaches have an unclear outcome. These are relatively rare in the overall dataset. Similar to the overall dataset, more attacks are successful than unsuccessful.

(e) **Password Guessing: Success and Strength Category** Only five scenes with password guessing and a shown password (allowing strength class analysis). Four times this is a class 0 password, one time a Class 1.



(d) **Password Hacking: Success** For the High-grossing movies, exactly half of the hacking attempts are successful and the other half have an unclear outcome. Not a single attempt fails. This is very different from the overall dataset, where most attacks are successful.

(f) **Password Hacking: Success and Strength Category** Not a single scene with password hacking and a shown password (allowing strength class analysis). Accordingly, no further analysis is possible.

Figure 12: Comparison of high-grossing movies and movie-dataset for password attacks. The number of codes per topic was calculated for each diagram for both the high-grossing and the overall dataset. This was used to calculate the percentages.