Geographic Hints for Passphrase Authentication

Alaadin Addas
Ontario Tech University
Oshawa, Canada
alaadin.addas@uoit.ca

Julie Thorpe
Ontario Tech University
Oshawa, Canada
julie.thorpe@uoit.ca

Amirali Salehi-Abari
Ontario Tech University
Oshawa, Canada
abari@uoit.ca

***Abstract*—We propose and study the use of geographic hints to aid memorability of passphrase-style authentication secrets. *Geographic hints* are map locations that are selected by the user at the time of passphrase creation, and shown to the user as a hint at the time of passphrase login. We implement the *GeoHints* system and analyze how geographic hints impact the usability and security of passphrase-style secrets in a multi-session user study (n=38). The study involved testing for multiple passphrase interference—each participant was asked to recall 4 distinct passphrases. Our study indicates that while geographic hints showed promise for reducing memory interference, GeoHints (as implemented) does not produce a viable authentication system, as the login success rate was 25% 7–11 days after passphrase selection. We analyze the root causes of login errors, finding that most were due to inexact recall of free-form text input. This finding points towards opportunities to improve the system design, and we suggest improvements that we believe will lead to viable systems that employ geographic hints.**

***Index Terms*—Authentication, Passphrases, Geographic Authentication, Hints, Geographic Hints, Secret Notes.**

I. Introduction

Passwords persist as popular primary authentication mechanisms despite their security flaws [1]–[3]. Password crackers have been effective in guessing large numbers of passwords [4]–[6]. Furthermore, the threat of password crackers has been exacerbated mainly due user password choice patterns that have been observed in leaked password datasets [4], [7], [8]. Although numerous measures are taken to make passwords more secure, they unfortunately have introduced some usability and security problems. Stringent password policies, for example, have largely been unsuccessful because they reduce memorability and increase the rate of input errors [9]. Passphrases, with the ambitious goal of improving both security and memorability, have failed due to input errors and memorability issues [10], [11]. As with Passwords, some secondary (or fallback) authentication mechanisms (e.g., security questions and personal knowledge questions) also suffer from similar security and usability flaws [3], [12]–[15].

Flaws in primary and secondary authentication mechanisms have motivated the investigation of alternative means of authentication including graphical passwords [16]–[19], implicit authentication [20] and geographic authentication [21]–[23]. The promising security and usability characteristics of geographical authentication (e.g., high memorability of 97% [22]) has motivated us to propose *GeoHints* where the secret is a passphrase-style text paired with a preset location on the map as a hint. GeoHints is a novel variant of GeoPass [22] and GeoPassNotes [23]. GeoPass authenticates users by setting a marker on a map whereas GeoPassNotes authenticates users by the combination of notes and locations. In contrast to GeoPass and GeoPassNotes, the location marker in GeoHints is a hint rather than being part of the secret. Our study design is also different than GeoPassNotes [23], and GeoPass [22], and includes a multiple password interference evaluation (see Section IV).

We investigate the usability, security, and multiple passphrase interference of GeoHints through a multi-session in-lab user study (n=38) involving four accounts, spanning two sessions over 7–11 days. Our results suggest that GeoHints provides security benefits over primary authentication mechanisms such as passwords, and secondary authentication mechanisms such as personal knowledge questions. The average character lengths of passphrases in GeoHints was 17 characters long, thus rendering it more resilient to guessing attacks. GeoHints offers increased resilience to throttled and unthrottled attacks when compared to passwords. Additionally, while GeoHints is not immune to classical phishing attacks, it is more resilient to classical phishing attacks than passwords.

In terms of usability, GeoHints shows promise for reducing memory interference between accounts. Only 7% of failed attempts were due to the interference effect in our study, which is an improvement when compared to that of passwords (78%) [24]. However, the login success rate of GeoHints was low, mainly due to inexact recall. Only 25% of login attempts were successful 7–11 days after setting the credentials. We analyze the root causes of participant errors, and present possible solutions to alleviate these issues. Our investigation has shed light on the lack of usability of long strings for authentication purposes, specifically in GeoHints. Future work should investigate the usability and security of GeoHints with selection-based inputs and/or shorter string inputs (while imposing some restrictions on the maximum string length).

II. Related Work

We begin by exploring literature related to the current primary and secondary authentication methods. We then discuss geographical authentication systems and variants of text-based authentication systems (passphrases). Lastly, we explore previous attempts that utilized hints to improve memorability.

Primary Authentication. Passwords and their variants are widely-used as primary authentication systems. However, passwords suffer from several security and usability problems.

Short and simple passwords are memorable but vulnerable to password cracking algorithms. On the other hand, long and complex passwords with a mixture of character types (e.g., numbers, symbols, lowercase letters, and uppercase letters) are more secure but difficult to remember [25]. Considerable attention is given to analyzing password choice patterns, and developing password guessers which exploit those patterns [4]–[8], [25].

Chou *et al.* [8], by analysing the RockYou leaked password dataset [26], found emerging patterns in passwords (e.g., lowercase letters followed by a number). Users also exhibit the tendency of reusing their passwords or simple variations of them over multiple platforms (e.g., Facebook, Hotmail, MySpace). It is estimated that 43%–51% of users utilize the same password across multiple platforms (e.g., Facebook, Hotmail, MySpace) [25]. Additionally, several key transformations are identified that users deploy to alter passwords across different platforms [25]. These discoveries facilitate the development of a powerful cross-site password guessing algorithm [25].

Veras *et al.* [4], [7] utilized natural language processing techniques to analyze semantic patterns in passwords. Leveraging several large-scale leaked password sets, they improved the performance of offline guessing attacks using probabilistic context free grammars on the LinkedIn and MySpace leaked passwords where 67% and 32% passwords were cracked, respectively, by only $2^{31.4}$ guesses. Durmuth *et al.* [5] successfully utilized Ordered Markov Enumerators (OME) for password cracking. Melicher *et al.* [6] successfully utilized Artificial Neural Networks (ANNs) for password cracking, and yielded state-of-the-art results. The related work highlights the security flaws of passwords as a means for primary authentication, motivating research into alternative means of authentication.

Secondary Authentication. The popular methods for secondary authentication are personal knowledge questions, email resets, and SMS resets. Vulnerabilities in secondary authentication mechanisms can render the primary method of authentication useless.

Just *et al.* analyzed answers to personal knowledge questions and found that they are easy to guess. Golla *et al.* [3] evaluated a leaked set of 3.9 million answers to personal knowledge questions, and reaffirmed the previous findings. Bonneau *et al.* [15] analyzed millions of password recovery attempts using personal knowledge questions and found that users failed 40% of the time.

Email resets are highly usable except where the user has lost access to the recovery email account [13]. It is noted that email resets make the email a single point of attack [13]. Guri *et al.* [14] found that rogue applications can request access to resources that compromise email resets and SMS resets [14]. As SMS resets rely on a secure channel of communication, flaws in telecommunication protocols [27], [28] can compromise the security of SMS rests.

Geographic Authentication. The GeoPass authentication system requires a user to select a point on a map as his/her secret [22]. Thorpe *et al.* [22] conducted a user study (n=35) on GeoPass spanning three sessions over 7 days. During the first session participants were asked to set and confirm their login credentials for GeoPass. Session 2 was held within 2 days of Session 1, 33 returned participants achieved a 100% login success rate given 5 attempts. Session 3 was scheduled one week after Session 1, 30 participants returned, and achieved a 97% login success rate within 5 attempts. GeoPassNotes [23], an extension of GeoPass, requires users to set a point on a map and associate a note with that location for authentication. A user study was conducted on GeoPassNotes with a similar design to that of GeoPass. In sessions 2 and 3, returned participants had 100% login success rate within 5 attempts.

Hang *et al.* [21] developed a geographical authentication system in lieu of traditional personal knowledge questions. Its users could answer either a predefined location question (e.g., where was your longest travel to?), a guided questions (e.g., please define a location-based question that refers to a travel/vacation destination), or a user-defined open-ended question. Users answer these questions by setting a marker on the map. A three-session user study (n=30) was conducted over 6 months. During Session 1, participants answered 3 predefined location questions, 3 guided questions, and 3 participant-defined open-ended questions. Session 2 was held 4 weeks later, and 90% of participants were able to recall their answers. Session 3 was held 6 months later and had a 92% login success rate for all returned participants.

Of the three reviewed geographical authentication systems, all exhibit very high memorability. However, GeoPassNotes [23] provides the most security benefits due to the pairing of geographical authentication and text based authentication (in the form of notes). Both GeoPass and Hang *et al.*'s system would require stringent system restrictions (e.g., stringent throttling, and blacklisting popular locations) in order to prevent online guessing attacks. The strong security and usability of GeoPassNotes motivated us to investigate the use of geographic hints with text-based authentication, we hypothesized that the addition of a geographic hint would increase memorability in the long term.

Passphrases. Passphrases relax stringent password policies (e.g., the requirement for lowercase letters, uppercase letters, numbers, etc.) in favor of a longer passphrase composed of different words [10]. Passphrases are intended to make text-based authentication more resilient to brute force attacks with their longer inputs [10]. It is shown that the memorability of passphrases is comparable to that of passwords with stringent policies [10]. It is also noted that input error was frequently resulting in a high number of failed attempts. While the security of passphrases might be slightly increased, the relaxation of the stringent password policies has not improved the usability of longer text inputs as passwords. Shay *et al.* [29] studied the usability of system assigned passphrases composed of 3 or 4 words, and compared them to system assigned passwords of similar entropy. They found that the usability of system-assigned passwords and passphrases is comparable

across different metrics. While both have similar memorability rates, system-assigned passphrases took longer to input and had frequent input errors.

Attempts to make passphrases more usable include relaxing error tolerance using string-edit distance metrics [30] and deploying implicit learning techniques (semantic priming and contextual cueing) [31]. Utilizing a string-edit distance decreases the number of failed attempts, hence increasing the overall usability of passphrases [30]. Implicit learning techniques have recently improved the memorability and login time of system assigned passphrases [31].

We attempt to leverage the security of longer text inputs and offset the memorability problems of passphrases through the utilization of a geographic hint. We also attempt to offset the input errors [10], [30] using a Levenshtein string edit distance.

Graphical Passwords.

Graphical passwords have been researched as an alternative to passwords [16]. A popular class of graphical passwords are click-based graphical passwords including PassPoints [17], Cued Click Points [19], Persuasive Cued Click Points [32], and PassPoints with the presentation effect [33]. These systems expect the users to choose and recall a sequence of points on a set of background images as their passwords. For usability issues, these systems authenticated users if their click-points have acceptable *error margin* to their selected points. Error margins are implemented through the process of discretization [18], [34], [35]. The security of various passpoint-style graphical passwords is studied [33], [36]–[39] which has motivated the development and design click-based authentication systems on videos [40], and digital maps [22], [23], [41].

Autobiographical Authentication.

Autobiographical authentication relies on data from day to day activities for authentication. Typically data from day to day activities is gathered from smartphone sensors [42]–[46]. The usability of several categories of autobiographical data for authentication is under question [42]–[46]. More specifically autobiographical location data for authentication has been investigated as an alternative means of fallback authentication, with varying degrees of success [43]–[46]. While autobiographical location data for authentication can be used in a manner that enhances system security, the usability is still very low when compared to other commonly utilized fallback authentication mechanisms (e.g., security questions, email resets, and SMS resets) [43]–[46].

Hint-Based Memorability Improvement. Hints proved to be effective in improving the memorability of graphical authentication schemes [47] and autobiographical authentication schemes [48]. In both instances, the hints were in the form of text, while GeoHints has a location-based hint.

III. GeoHints: System Design

GeoHints, developed with the Google Maps API [49], allows users to set a location and an associated passphrase-style secret note to the location. The location is utilized as a hint and the user is not required to remember it, whereas the secret note is required for authentication. Users can get to a location by searching or dragging on the map. Ideally, users will select locations that have an association with the secret note to aid in memorability, without labelling the location. Users can also zoom in and out. Our system stores both the location and the associated secret note in a database. See Fig. 1 for an interface screenshot.

The text-area shown in Fig. 1 was deliberately enlarged to nudge the users to input longer secret notes for authentication. As this could cause input and memorability errors which impact the rate of successful authentication, GeoHints allows a user to successfully authenticate if the inputted secret note has at least Levenshtein distance of 0.8 to the selected secret note.[1] Users are also asked to confirm the secret note after setting it. For confirmation, there must be an exact match. This is consistent with current authentication systems that ask the user to confirm their password during the registration to improve memorability and to ensure that no input error has occurred.

We reiterate that GeoHints is a novel variant of GeoPass [22], and GeoPassNotes [23]. GeoHints was tested utilizing a different study design with multiple passwords interference, and relies on a different type of secret than GeoPassNotes and GeoPass.

IV. User Study

We evaluate the security and usability of GeoHints through an in-lab user study (n=38, 19 pairs), approved by our university's Research Ethics Board. Prior to the user study, GeoHints was pilot tested by 8 volunteers including colleagues, friends, and family members. 4 pilot testers were experienced computer users with degrees in Computer Science or IT while others were causal computer users. The pilot testing allowed us to debug our system, and improve the user study's instructions.

For the user study, the participants were recruited by posters on campus and a broadcast email. Participation was limited to students, visitors, and employees of our university. All participants must have met the following criteria: (i) At least 18 years old; (ii) Participants must bring a pair. Our user study contained two sessions spanning over 7-11 days. The pairs completed the exact same steps (i.e., we did not have a main participant within a pair).

Session 1. This session was conducted in-lab with multiple different time-slots. A maximum of 2 pairs (4 participants) were allowed in any time slot. Each participant was awarded $\$8$ for their participation. At the beginning of the session, participants were asked to sit across from each other to avoid any contamination of results. Afterwards, we read a set of pre-written instructions and ran a demo of GeoHints for the participants. They were then asked to use GeoHints by their laptop. They were first required to agree to the consent form, if they wished to continue. Next, they answered

[1]The Levenshtein distance was converted into a real number representing the percentage of similarity with regard to the length of the target string. For example, if the Levenshtein distance is 0.8 with a target string length of 5 characters, that means there was one substitution, deletion, or insertion.

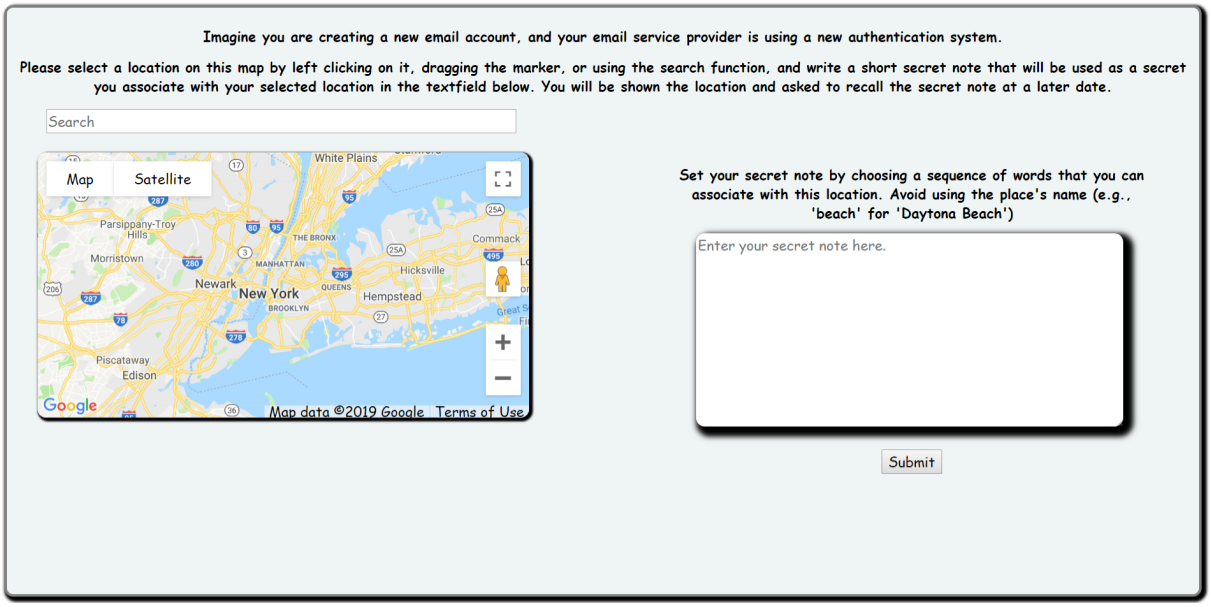

Fig. 1: GeoHints Interface

some demographic questions (e.g., age, gender, academic background, and relationship to their pairs).To test GeoHints for the multiple password interference, we ask users to create GeoHints credentials for four hypothetical accounts: an email account, a bank account, an e-commerce account, and a social media account.

Session 2. Session 2 was conducted 7–11 days after Session 1. 17 pairs returned for Session 2. The participants were compensated $10, and if a pair completed Session 1 and Session 2 in full, they were entered into a draw for $100. After reading the instructions and demonstrating the GeoHints system, participants were asked to recall their four GeoHints credentials set $7-11$ days earlier. Participants were given a maximum of five login attempts. To test the resilience of the system to guessing attacks, our participants were asked to guess their pair's secret note given the location as a hint. During the guessing phase of the study, participants were actively encouraged to utilize the internet for research, if needed. Lastly, we had an exit survey with several usability questions.

In our study, we also tested another completely independent authentication system [45], [46] that is not discussed in this paper. In both sessions, the GeoHints memory test was performed before testing the other independent system.

Demographics Details. Recruited participants were all undergraduate students in the range of 18–30 years old (average = 21.3). Out of 38 initial participants, 13 were female (34.2%), 25 were male (65.7%), and 15 (39%) had already taken some computer security/IT course .

V. Results

We evaluate GeoHints for security and usability using key metrics relevant to a text-based authentication system with a geographical hint.

A. Security Analysis

GeoHints was evaluated for its resilience to throttled guessing attacks, unthrottled guessing attacks, and classical phishing attacks.

Throttled Guessing Attacks. Following Bonneau *et al.* [2], we consider a system to be resilient to throttled guessing attacks if it withstands an attacker with the capability of making 10 guesses a day for 365 days (3650 guesses a year) and not compromise more than 1% of accounts [2]. We tested GeoHints for two classes of throttled online attacks where the classes differentiate by whether or not the attacker has the first-hand knowledge of the potential victim. The attacker with the first-hand knowledge of potential victim is referred to as *known adversary* [50].

We tested against throttled known adversary online attackers by asking the participating pairs to guess each others' secret notes. Fig. 2 showcases the results in terms of false positive rate (FPR) and true positive rate (TPR). With our threshold of 0.8 Levenshtein distance for a successful authentication, only 2.9% (4/136) secret notes were guessed correctly by our adversarial pairs. Therefore, while GeoHints is not considered resilient to known adversary throttled online guessing attacks, it is still highly immune to adversaries with first-hand knowledge.

We test against throttled online attacks (without first-hand knowledge of potential victim) by utilizing three password cracking algorithms: i) John the Ripper in incremental mode [51]; ii) a probabilistic context free grammar [52]; and iii) a semantic cracker [4]. The most successful password cracker was the semantic cracker [4]. A total of 22/136 (16.1%) secret

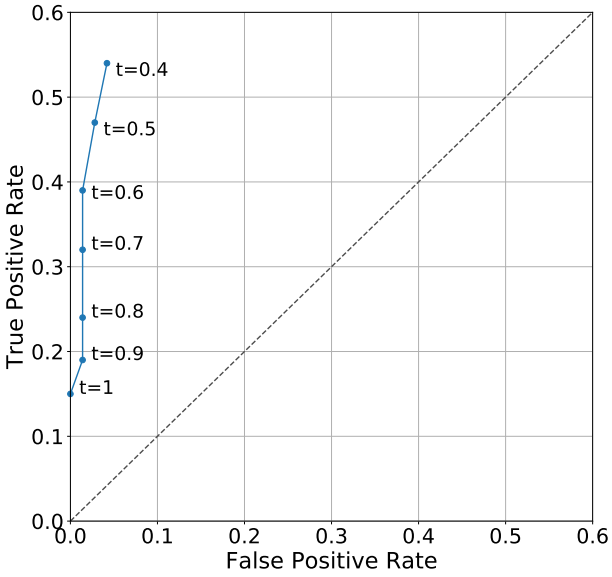

Fig. 2: Receiver operating characteristic curve: the TPR vs. FPR with various Levenshtein distance thresholds.

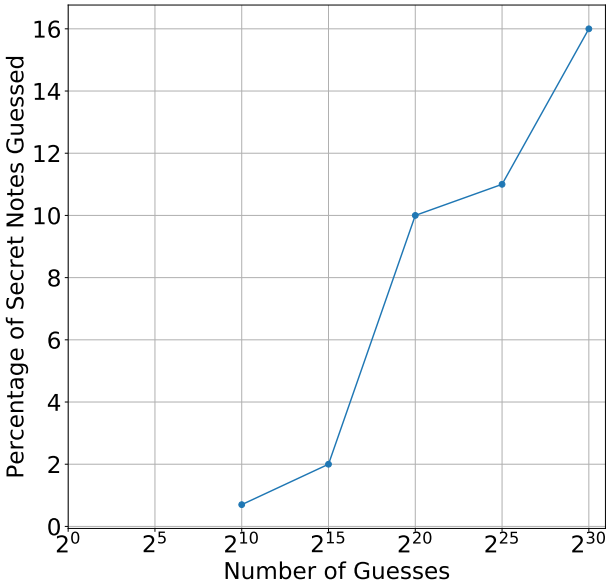

Fig. 3: The percentage of all secret notes of four hypothetical accounts guessed correctly by the semantic cracker.

notes were successfully guessed by the semantic cracker [2] (see Fig. 3). At 3650 ($2^{11.8}$) guesses the semantic cracker failed to guess more than 1% of the secret notes. Hence, GeoHints is resilient to throttled online guessing attacks with no first-hand knowledge.

Unthrottled Guessing Attacks. An authentication system is resilient to unthrottled guessing attacks if no more than 1% of accounts are compromised with 2^{30} guesses [2]. As shown in Fig. 3, GeoHints is not resilient to unthrottled guessing attacks since 16.1% of secret notes were guessed by 2^{30} guesses. However, it is more immune than typical passwords as the notes were relatively long with an average of 17 characters.

Classical Phishing Attacks. In a classical phishing attack, the attacker mimics a legitimate login portal for the purpose of tricking the user into entering their authentication credentials[3]. For password authentication, the credentials are typically a username and password (both are provided by the user). However in the case of GeoHints, the geographic hint must be given to the user. This feature does not make GeoHints immune to classical phishing attacks, but it makes GeoHints more resistant to classical phishing attacks than password-style primary authentication methods.

B. Usability Analysis

We evaluate the usability of GeoHints using three key metrics: (i) login success rate, (ii) login time, and (iii) credential creation time.

Login Success Rate. The login success rate of GeoHints is 25% 7–11 days after setting the credentials. The TPR is significantly improved to 36% when the Levenshtein distance is set to 0.6 while the FPR remains at 1.4% (see Fig. 2). However, setting the Levenshtein distance to 0.6 can compromise the security of GeoHints to a great extent when deployed in practice. Fig. 4 shows the correct/incorrect successful authentication (within 5 attempts) for all participants of each account. The low login success rate of GeoHints makes it less suitable as a candidate for the replacement of current primary or secondary authentication systems. We discuss our qualitative analysis of the reasons behind this low success rate in Sec. VI.

Login Time. GeoHints is prone to user input error despite using Levenshtein's distance of 0.8. Participants had an average of 3.26 (std= ± 0.52) failed attempts until a successful login. Only 9% (3/34) of participants successfully logged in to all four accounts[4]. The high rate of failure (before a successful login) made the average login time 100 seconds. Fig. 5 shows the average login time (including failures) for all four hypothetical accounts.

Credential Setting Time. The required time for creating a credential is an important usability factor, which can contribute to a user's abandonment, or cognitive fatigue. GeoHints had an average credential setup time of 2.39 minutes (credential creation + confirmation) per account. Frequent confirmation errors led to the high credential setup time. Fig. 6 shows the average credential setup time for each of the four hypothetical accounts (in the order they were presented). The average

[2]When testing the resilience of GeoHints against different password cracking algorithms a Levenshtein distance of 0.8 was not used. Exact matches were required for a guess to be considered successful. This might have resulted in slight overestimate of the security of GeoHints.

[3]In a classical phishing attack, we assume that the attacker does not have access to a username, and we assume that the login portal is not customized for a particular potential victim.

[4]Of the three participants that successfully logged in to all four hypothetical accounts, one participant had the same secret note for all 4 accounts.

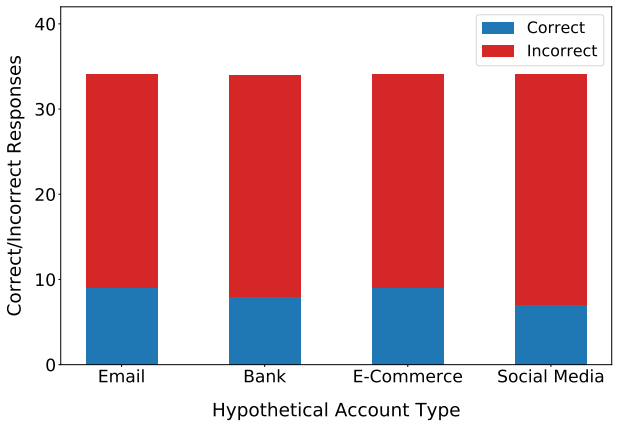

Fig. 4: Comparison of correct and incorrect responses for all users across all four different account types (n=34), given 5 attempts.

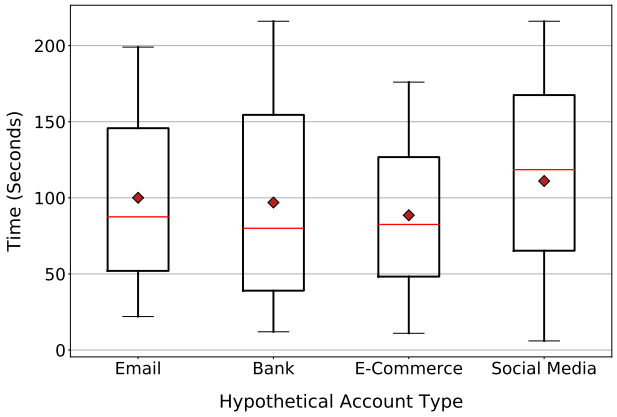

Fig. 5: Login time including failures (n=34).

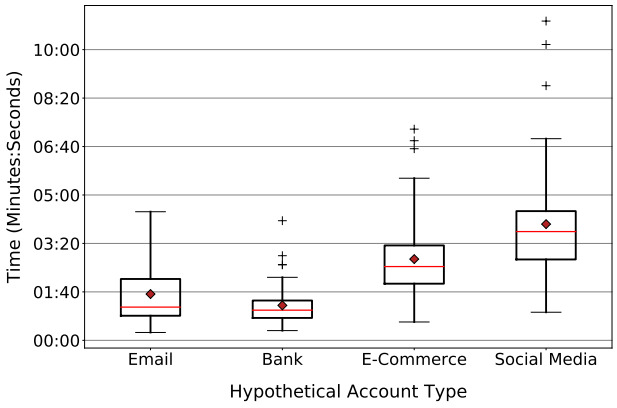

Fig. 6: Credential creation time for various account types.

credential setup time is higher than that of popular primary and secondary authentication systems.

VI. Discussion and Future Work

GeoHints provides some security advantages over passwords. GeoHints offers increased resilience to throttled online guessing attacks, and to unthrottled guessing attacks. When emulating a throttled guessing attack given $2^{11.8}$ guesses [2], the password crackers we used [4], [51], [52] failed to compromise more than 1% of accounts. When we emulated an unthrottled guessing attack given 2^{30} guesses, the most successful password cracker [4] guessed 16.1% of the secret notes. This is a significant improvement because the same password cracker guessed 67% of the LinkedIn leaked password dataset. Furthermore, GeoHints also has slightly enhanced resilience to phishing when compared to primary authentication mechanisms (e.g., passwords) and secondary authentication mechanisms such as personal knowledge questions. However, the usability of GeoHints in terms of login success rate, login time, and credential setup time is not comparable to popular primary and secondary authentication mechanisms. Reasons for the low true positive rate of 25% (7–11 days after credential setup) are explored in this section. Our analysis has shown that only 36.2% of all failed attempts were due to not remembering a secret note (see Table I).

We coded all the failed attempts into three general categories: (i) *inexact recall*, (ii) *memory loss*, and (iii) *interference*.[5] See Table I for details. Inexact recall is a parent group which encompasses *semantic similarity* and *rewording*. Semantically similar login attempts are similar in meaning to the set secret note but incorrect because it was not at 0.8 Levenshtein distance to the set secret note (e.g., the set note was "this is my home" while the recall attempt was something relating to the home). Rewording includes those failed login attempts that were similar to the originally set secret note but were reworded in some way that led to a failure (e.g., "I saw the raptors play here" as the set secret note and the attempt was "This is where I saw the raptors play"). Both semantic similarity and rewording fall into the inexact recall group because it shows a pattern of the geographic hint aiding the participants in recalling the topic of their secret note. Interference refers to the effect of having multiple accounts utilizing the same authentication method, where users can get mixed up regarding which secret note belongs to an account. Our classification of failed attempts provides an explanation to why the true positive rate was low: the main cause was inexact recall (56.1% of login failures). Login failures were coded as complete memory loss when they were not detected as inexact recall or interference; only 36.2% of failed attempts were coded as such. This suggests that the geographic hints were triggering the participants' memory regarding the overarching topic, but the users were making mistakes in the sentence structure (e.g., "This is my home" as a preset secret note, with a login attempt of" "my house" which constitutes a failure).

The long length of a secret note (average = 17 characters long) is due to participants typing in full sentences related to the location hint. This was partly because of the design of GeoHints, where we input a large text area, thus nudging the participants to a long passphrase-style secret note. The actual minimum length was only 5 characters. We draw a

[5]Authors manually and independently coded all the failed attempts.

Category	Number of Failed Attempts
Inexact Recall	155 (56.1%)
Memory Loss	100 (36.2%)
Interference	21 (7.6%)

TABLE I: Categorization of failed attempts.

Category	Number of Notes
Direct Label	51 (37.5%)
Descriptive Label	35 (25.7%)
Direct Naming	5 (3.6%)
Other	41 (30.1%)

TABLE II: Breakdown of secret notes into categories.

comparison between this effect and leaving extra space to answer a question on an exam, where extra space prompts a student to write more even if the question can be satisfied with a one sentence answer. We believe full sentences have contributed to the high chance of inexact recall errors.

The interference effect has accounted for only 21 (7.6%) failed attempts. Most of those interference failed attempts ended up with a successful login or an inexact recall error which then led to a complete failure (not a complete memory loss). The low rate of multiple password interference was due to the geographic hint according to our analysis. When compared to text-based passwords (78% failure due to interference [24]), and GeoPass (41.5% failure due to interference [53]), GeoHints (7.6% failure due to interference) has a lower rate of failed login attempts due to interference effects.

Because hints in GeoHints attempt to trigger the user's memory, it can also provide an attacker with contextual information to successfully guess a secret note. We prompt users to not directly label locations in the creation phase using the following prompt: "Set your secret note by choosing a sequence of words that you can associate with this location. Avoid using the place's name (e.g., "beach" for "Daytona Beach")". This prompt is borrowed from the GeoPassNotes study [23]. Despite our warning prompt, our secret note analysis revealed labelling.

We coded the secret notes into four categories based on the relationship between the geographic hint and the secret note (see Table II).[6]. The categories are *direct label* (e.g., this is my home), *descriptive label* (e.g., a place I won a basketball game), *direct naming* (e.g., directly naming a restaurant), and *other* (e.g., not a label of any kind). The first three categories (direct label, descriptive label, and direct naming) are all types of labels, Table II shows the breakdown of all secret note categorizations according to our analysis.

We evaluated the impact of labelling using our results from the pair guessing phase of our study. Since only $4/136$ (2.9%) of secret notes were compromised (three of which were labels), our preliminary assessment suggests that labelling did not have a grave security impact. However, given a more determined adversary this trend of labelling could become a problem should GeoHints be deployed in practice. Blacklisting certain landmarks and banning labels using system enforced rules is a potential solution to the labelling problem. Labelling did however have a positive impact on usability. Labelled secret notes were more likely to result in a successful login attempt, $27/34$ (79.4%) of the secret notes that were successfully recalled were labels (this includes labels, descriptive labels, and direct naming), while only $7/34$ (20.5%) of secret notes that were not labels were successfully recalled.

Our analysis of failed attempts and the secret notes inform the future implementations of variants of GeoHints or any systems which rely on geographic hints to aid in the memorability of a passphrase-style secret note. In order to avoid the usability flaws that we encountered, the following guidelines we extracted based on our analysis should be applied:

1) **Use selection-based input.**

   Our analysis of failed attempts revealed that the majority of login failures were due to inexact recall. Inexact recall issues can be eliminated through the use of a selection-based interface, whereby users must recognize and select the input from a set of alternatives. One example of such an interface was employed for system-assigned passphrases [31]. While we believe this is a promising approach for future interfaces, it does present a security and usability trade-off, and as such we also present recommendations for free-form passphrase interfaces.

2) **Set a policy for white-space and punctuation.**

   During the study we noticed that some participants were confused whether they should include white-space and punctuation because they usually do not do so with passwords (but it is the norm with natural sentences). Secret notes are incredibly similar to passwords which causes that confusion. Therefore, it would be helpful to include a note while the users are setting their credentials that white-spaces and punctuation are okay.

3) **Encourage association.**

   As mentioned earlier, labelling of any kind does improve the chance that a secret note will be successfully remembered 7–11 days after it was set. While direct labelling and direct naming should be discouraged for fear of a security compromise, it is important to associate the secret note and the selected location with a memory of some sort. Many labels analyzed were a type of association, and that yielded a higher success rate. Therefore, changing the instructions to include an emphasis on association with some sort of memory could enhance the memorability of the secret note to a great extent. Al-Ameen *et al.* [53] utilized the mental story approach for improving the memorability of GeoPass and yielded positive memorability results; the same approach can be utilized for GeoHints.

4) **Set and confirm on the same page.**

   In our system design of GeoHints, the participant had to set the location and the secret note. After they clicked

[6]Authors manually analyzed and independently coded the secret notes.

submit, the secret note had to be confirmed. However, in lieu of that configuration, we propose that the secret note be set, and secret note confirmation occur on the same page much like how passwords are set and confirmed on the same page. We make this recommendation in an effort to decrease the time required to set GeoHints credentials which would in turn make it more usable. We observed many participants having difficulty in the confirmation phase, making the credential setup time longer.

Future work on GeoHints (or its variants) are required to improve its usability. Our results suggest that using a large text area for nudging participants for a long secret is effective for improving security, but a major culprit to usability. Our motivation behind a large text-area was to ensure a secret note with enough entropy to withstand an unthrottled guessing attack. However, the downgrade in usability was exemplified in the high failure rates. Future work can investigate the optimal nudging for secret note's length for balancing security strength and usability.

VII. Limitations

One core limitation of our work is that we only tested GeoHints in a desktop browser environment, future work can explore the security and usability of GeoHints in a mobile browser environment as human factors in authentication play a role depending on the target environment [54]. It is possible that the average length of a secret note to be significantly different on a mobile device, and consequently the usability and security metrics may change.

One limitation of our experiment is that we did not utilize a Levenshtein distance of 0.8 when testing the resilience of GeoHints against different password cracking algorithms.

While we test GeoHints's resilience to known adversary attacks by recruiting study participants in pairs and asking them to guess each other's secret notes, a further analysis is required to investigate the potency of known adversary attacks in light of semantic preferences (e.g., frequently labelling a home or workplace) that are evident in the secret notes. This may have affected our online throttled attacks results, in light of the amount of publicly available information on social media platforms.

VIII. Conclusion

We designed and implemented GeoHints, an authentication system in which users are required to recall a passphrase-style secret note while a geographic hint is set to help the recall. We tested GeoHints for usability, security, and multiple passphrase interference through a multi-session in-lab user study spanning over 7–11 days. The results showed that GeoHints does offer certain security benefits (e.g., increased resilience to throttled and unthrottled guessing attacks) over primary authentication methods (e.g., passwords) and secondary authentication methods (e.g., personal knowledge questions). While the usability of the present implementation is relatively low when compared to the aforementioned primary and secondary authentication methods, our analysis showed that geographic hints show promise in reducing memory interference. There are also promising approaches that can be employed to increase the login success rate by dealing with the problem of inexact recall. Our study and analysis of GeoHints paves the way for interesting future work and investigation on the use of geographic hints.

IX. Acknowledgment

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference numbers 402500-2013 and RGPIN-2018-05903.

References

[1] J. Bonneau and S. Preibusch, "The password thicket: Technical and market failures in human authentication on the web," in *proceedings of the 9th Workshop on the Economics of Information Security (WEIS'10)*, 2010.

[2] J. Bonneau, C. Herley, P. C. Van Oorschot, and F. Stajano, "The quest to replace passwords: A framework for comparative evaluation of web authentication schemes," in *Proceedings of the 2012 IEEE Symposium on Security and Privacy (IEEE S&P'12)*, 2012, pp. 553–567.

[3] M. Golla and M. Dürmuth, "Analyzing 4 million real-world personal knowledge questions (short paper)," in *Proceedings of the 9th International Conference on Passwords*, 2015, pp. 39–44.

[4] R. Veras, C. Collins, and J. Thorpe, "On semantic patterns of passwords and their security impact," in *proceedings of the 2014 Network and Distributed System Security Symposium (NDSS'14)*, 2014.

[5] M. Dürmuth, F. Angelstorf, C. Castelluccia, D. Perito, and C. Abdelberi, "Omen: Faster password guessing using an ordered markov enumerator," in *Proceedings of the 2015 International Symposium on Engineering Secure Software and Systems (ESSoS'15)*, 2015.

[6] W. Melicher, B. Ur, S. M. Segreti, S. Komanduri, L. Bauer, N. Christin, and L. F. Cranor, "Fast, lean, and accurate: Modeling password guessability using neural networks," in *Proceedings of the 25th USENIX Security Symposium (USENIX'16)*, 2016, pp. 175–191.

[7] R. Veras, J. Thorpe, and C. Collins, "Visualizing semantics in passwords: The role of dates," in *Proceedings of the Ninth International Symposium on Visualization for Cyber Security (VizSec'12)*, 2012, pp. 88–95.

[8] H.-C. Chou, H.-C. Lee, H.-J. Yu, F.-P. Lai, K.-H. Huang, and C.-W. Hsueh, "Password cracking based on learned patterns from disclosed passwords," *International Journal of Inovative Computing, Information, and Control (IJICIC)*, vol. 9, pp. 821–839, 2013.

[9] S. Komanduri, R. Shay, P. G. Kelley, M. L. Mazurek, L. Bauer, N. Christin, L. F. Cranor, and S. Egelman, "Of passwords and people: Measuring the effect of password-composition policies," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11)*, 2011, pp. 2595–2604.

[10] M. Keith, B. Shao, and P. J. Steinbart, "The usability of passphrases for authentication: An empirical field study," *International Journal of Human-Computer Studies*, vol. 65, pp. 17–28, 2007.

[11] R. Shay, P. G. Kelley, S. Komanduri, M. L. Mazurek, B. Ur, T. Vidas, L. Bauer, N. Christin, and L. F. Cranor, "Correct horse battery staple: Exploring the usability of system-assigned passphrases," in *Proceedings of the 8th Symposium on Usable Privacy and Security (SOUPS'12)*, 2012, pp. 7:1–7:20.

[12] M. Just and D. Aspinall, "Personal choice and challenge questions: A security and usability assessment," in *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS'09)*, 2009, pp. 8:1 – 8:11.

[13] S. L. Garfinkel, "Email-based identification and authentication: An alternative to pki?" *IEEE Security & Privacy*, vol. 99, pp. 20–26, 2003.

[14] M. Guri, E. Shemer, D. Shirtz, and Y. Elovici, "Personal information leakage during password recovery of internet services," in *Proceedings of the 2016 European Intelligence and Security Informatics Conference (EISIC'16)*, 2016, pp. 136–139.

[15] J. Bonneau, E. Bursztein, I. Caron, R. Jackson, and M. Williamson, "Secrets, lies, and account recovery: Lessons from the use of personal knowledge questions at google," in *Proceedings of the 24th International Conference on World Wide Web (WWW'15)*, 2015, pp. 141–150.

[16] R. Biddle, S. Chiasson, and P. Van Oorschot, "Graphical passwords: Learning from the first twelve years," *ACM Computing Surveys (CSUR)*, vol. 44, pp. 19:1–19:41, 2012.

[17] S. Wiedenbeck, J. Waters, J.-C. Birget, A. Brodskiy, and N. Memon, "Passpoints: Design and longitudinal evaluation of a graphical password system," *International Journal of Human-Computer Studies*, vol. 63, pp. 102–127, 2005.

[18] X. Suo, Y. Zhu, and G. S. Owen, "Graphical passwords: A survey," in *21st Annual Computer Security Applications Conference (ACSAC'05)*, 2005, pp. 472–481.

[19] S. Chiasson, P. C. Van Oorschot, and R. Biddle, "Graphical password authentication using cued click points," in *European Symposium on Research in Computer Security (ESORICS'07)*, 2007, pp. 359–374.

[20] H. Khan, U. Hengartner, and D. Vogel, "Usability and security perceptions of implicit authentication: Convenient, secure, sometimes annoying," in *Eleventh Symposium On Usable Privacy and Security (SOUPS'15)*, 2015, pp. 225–239.

[21] A. Hang, A. De Luca, M. Smith, M. Richter, and H. Hussmann, "Where have you been? using location-based security questions for fallback authentication," in *Proceedings of the 11th Symposium On Usable Privacy and Security (SOUPS'15)*, 2015, pp. 169–183.

[22] J. Thorpe, B. MacRae, and A. Salehi-Abari, "Usability and security evaluation of geopass: A geographic location-password scheme," in *Proceedings of the 9th Symposium on Usable Privacy and Security (SOUPS'13)*, 2013, pp. 14:1–14:14.

[23] B. MacRae, A. Salehi-Abari, and J. Thorpe, "An exploration of geographic authentication schemes," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 1997–2012, 2016.

[24] S. Chiasson, A. Forget, E. Stobert, P. C. van Oorschot, and R. Biddle, "Multiple password interference in text passwords and click-based graphical passwords," in *Proceedings of the 16th ACM Conference on Computer and Communications Security (CCS'09)*, 2009, pp. 500–511.

[25] A. Das, J. Bonneau, M. Caesar, N. Borisov, and X. Wang, "The tangled web of password reuse," in *Network and Distributed System Security Symposium (NDSS'14)*, 2014, pp. 23–26.

[26] "Rockyou passwords, https://wiki.skullsecurity.org/Passwords," site accessed June 2018.

[27] B. Welch, "Exploiting the weaknesses of SS7," *Network Security*, vol. 2017, pp. 17–19, 2017.

[28] "SS7 hack tutorial, https://fedotov.co/ss7-hack-tutorial-software-video," site accessed June 2018.

[29] R. Shay, P. G. Kelley, S. Komanduri, M. L. Mazurek, B. Ur, T. Vidas, L. Bauer, N. Christin, and L. F. Cranor, "Correct horse battery staple: Exploring the usability of system-assigned passphrases," in *Proceedings of the Eighth Symposium on Usable Privacy and Security (SOUPS'12)*, 2012, pp. 7:1–7:20.

[30] G. V. Bard, "Spelling-error tolerant, order-independent pass-phrases via the damerau-levenshtein string-edit distance metric," in *Proceedings of the Fifth Australasian Symposium on ACSW Frontiers (ACSW '07)*, 2007, pp. 117–124.

[31] Z. Joudaki, J. Thorpe, and M. V. Martin, "Reinforcing system-assigned passphrases through implicit learning," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS'18)*, 2018, pp. 1533–1548.

[32] S. Chiasson, A. Forget, R. Biddle, and P. C. Van Oorschot, "Influencing users towards better passwords: Persuasive cued click-points," in *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction*, 2008, pp. 121–130.

[33] J. Thorpe, M. Al-Badawi, B. MacRae, and A. Salehi-Abari, "The presentation effect on graphical passwords," in *Proceedings of the 2014 SIGCHI Conference on Human Factors in Computing Systems (CHI'14)*, 2014, pp. 2947–2950.

[34] J.-C. Birget, D. Hong, and N. Memon, "Robust discretization, with an application to graphical passwords," *IACR Cryptology ePrint Archive*, vol. 2003, pp. 168–177, 2003.

[35] S. Chiasson, J. Srinivasan, R. Biddle, and P. C. van Oorschot, "Centered discretization with application to graphical passwords (full paper)," in *Proceedings of the 1st Conference on Usability, Psychology, and Security (UPSEC'08)*, 2008, pp. 6:1–6:9.

[36] A. Salehi-Abari, J. Thorpe, and P. C. v. Oorschot, "On purely automated attacks and click-based graphical passwords," in *Proceedings of the 2008 Annual Computer Security Applications Conference (ACSAC'08)*, 2008, pp. 111–120.

[37] P. C. van Oorschot, A. Salehi-Abari, and J. Thorpe, "Purely automated attacks on passpoints-style graphical passwords," *IEEE Transactions on Information Forensics and Security*, vol. 5, pp. 393–405, 2010.

[38] J. Thorpe and P. C. van Oorschot, "Human-seeded attacks and exploiting hot-spots in graphical passwords," in *Proceedings of 16th USENIX Security Symposium (SS'07)*, 2007, pp. 8:1–8:16.

[39] Z. Zhao, G.-J. Ahn, and H. Hu, "Picture gesture authentication: Empirical analysis, automated attacks, and scheme evaluation," *ACM Transactions on Information and System Security (TISSEC)*, vol. 17, p. 14, 2015.

[40] J. Thorpe, A. Salehi-Abari, and R. Burden, "Video-passwords: Advertising while authenticating," in *Proceedings of the 2012 New Security Paradigms Workshop (NSPW'12)*, 2012, pp. 127–140.

[41] A. Hang, A. De Luca, E. Von Zezschwitz, M. Demmler, and H. Hussmann, "Locked your phone? buy a new one? from tales of fallback authentication on smartphones to actual concepts," in *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI'15)*, 2015, pp. 295–305.

[42] A. Hang, A. De Luca, and H. Hussmann, "I know what you did last week! do you?: Dynamic security questions for fallback authentication on smartphones," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI'15)*, 2015, pp. 1383–1392.

[43] Y. Albayram and M. M. H. Khan, "Evaluating smartphone-based dynamic security questions for fallback authentication: A field study," *Human-Centric Computing and Information Sciences*, vol. 6, p. 16, 2016.

[44] S. Das, E. Hayashi, and J. I. Hong, "Exploring capturable everyday memory for autobiographical authentication," in *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'13)*, 2013, pp. 211–220.

[45] A. Addas, A. Salehi-Abari, and J. Thorpe, "Geographical security questions for fallback authentication," in *Proceedings of the 17th Annual Conference on Privacy, Security, and Trust (PST'19)*, pp. 1–7.

[46] A. Addas, J. Thorpe, and A. Salehi-Abari, "Geographical security questions for fallback authentication," *CoRR*, vol. arXiv:1907.00998v1, pp. 1–18, 2019.

[47] S. Chowdhury, R. Poet, and L. Mackenzie, "Passhint: Memorable and secure authentication," in *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems (CHI'14)*, 2014, pp. 2917–2926.

[48] Y. Albayram and M. M. H. Khan, "Evaluating the effectiveness of using hints for autobiographical authentication: A field study," in *proceedings of the 11th Symposium on Usable Privacy and Security (SOUPS'15)*, 2015, pp. 211–224.

[49] "Google Maps API, https://developers.google.com/maps/documentation/javascript/tutorial," site accessed June 2018.

[50] A. Addas, J. Thorpe, and A. Salehi-Abari, "Towards models for quantifying the known adversary," in *Proceedings of the 2019 Workshop on New Security Paradigms (NSPW'19)*, 2019.

[51] "John the ripper password cracker, http://www.openwall.com/john/," site accessed June 2018.

[52] M. Weir, S. Aggarwal, B. De Medeiros, and B. Glodek, "Password cracking using probabilistic context-free grammars," in *Proceedings of the 30th IEEE Symposium on Security and Privacy (IEEE S&P'09)*, 2009, pp. 391–405.

[53] M. N. Al-Ameen and M. K. Wright, "A comprehensive study of the geopass user authentication scheme," *CoRR*, vol. abs/1408.2852, p. 6, 2014.

[54] W. Melicher, D. Kurilova, S. M. Segreti, P. Kalvani, R. Shay, B. Ur, L. Bauer, N. Christin, L. F. Cranor, and M. L. Mazurek, "Usability and security of text passwords on mobile devices," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI'15)*, 2016, pp. 527–539.