






Is It Really You Who Forgot the Password? When Account Recovery Meets Risk-Based Authentication

Andre Büttner¹✉, Andreas Thue Pedersen¹, Stephan Wiefling²,
Nils Gruschka¹, and Luigi Lo Iacono³

¹ University of Oslo, Oslo, Norway

{andrbut, nilsgrus}@ifi.uio.no

² swiefling.de, Sankt Augustin, Germany

ubisec23@swiefling.de

³ H-BRS University of Applied Sciences, Sankt Augustin, Germany

luigi.lo_iacono@h-brs.de

Abstract. Risk-based authentication (RBA) is used in online services to protect user accounts from unauthorized takeover. RBA commonly uses contextual features that indicate a suspicious login attempt when the characteristic attributes of the login context deviate from known and thus expected values. Previous research on RBA and anomaly detection in authentication has mainly focused on the login process. However, recent attacks have revealed vulnerabilities in other parts of the authentication process, specifically in the account recovery function. Consequently, to ensure comprehensive authentication security, the use of anomaly detection in the context of account recovery must also be investigated.

This paper presents the first study to investigate risk-based account recovery (RBAR) in the wild. We analyzed the adoption of RBAR by five prominent online services (that are known to use RBA). Our findings confirm the use of RBAR at Google, LinkedIn, and Amazon. Furthermore, we provide insights into the different RBAR mechanisms of these services and explore the impact of multi-factor authentication on them. Based on our findings, we create a first maturity model for RBAR challenges. The goal of our work is to help developers, administrators, and policy-makers gain an initial understanding of RBAR and to encourage further research in this direction.

Keywords: Risk-Based Account Recovery · RBAR · Authentication · Account Security · Online Services.

1 Introduction

Passwords are still the pre-dominant authentication method for online services, even for services that give access to confidential data or financial resources [14, 31]. However, attacks on password authentication can be automated—e.g., credential stuffing using leaked passwords—and therefore scaled with little effort. This makes account takeover attacks on password-protected online services very

Paper published at UbiSec '23.

This version of the contribution has been accepted for publication, after peer review (when applicable) but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections.

The Version of Record is available online at: http://dx.doi.org/10.1007/978-981-97-1274-8_26.

Use of this Accepted Version is subject to the publisher's Accepted Manuscript terms of use <https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms>

lucrative for hackers [3]. As a countermeasure, more and more services offer multi-factor authentication (MFA) as an extension to password authentication. In this case, the user has to give additional proof of their identity, e.g., by entering a code from a one-time password (OTP) app or a text message (SMS). However, the additional step makes the authentication process more cumbersome and increases the risk of account lockouts in case the additional token gets lost [30].

The idea of risk-based authentication (RBA) [12, 14, 38] is to balance security and usability. Here, the online service only requests additional authentication steps or blocks a client when it detects suspicious behavior. RBA does this by analyzing a set of feature values (e.g., location, browser, or login time) during the login process [14, 38].

A general problem with authentication is that the user might lose access to the authentication method—in the case of password authentication, this means primarily forgetting the password. In such a case, the user has to pass the *account recovery* process to regain access to their account. The process often involves sending a password reset link or an OTP to a pre-configured email address or phone number. If the required authentication (e.g., ownership of a phone, login to the email account) is weaker than the primary authentication, account recovery puts the overall account security at risk [27, 29].

A high and common threat to account recovery mechanisms via email is when an attacker gains access to the corresponding email account, e.g., via credential stuffing [2, 33]. The recent FBI cybercrime report [11] shows that compromised email addresses and phishing attacks are very popular attacks with potentially high financial loss for the hacked victims. Therefore, it is very important for online services to secure account recovery, for example, with MFA or RBA. So far, risk-based mechanisms have mostly been studied in the context of login authentication. However, we observed that mechanisms similar to RBA are also used for account recovery.

We define *Risk-Based Account Recovery* (RBAR)⁴ as a dynamic account recovery process on online services. It was indicated that such a method is used at a large online service [7], but beyond that, RBAR and its appearances in the wild have not been publicly investigated yet. This is, however, important as it has the potential to protect a large number of users from account recovery attacks immediately. To learn about the current use of RBAR, we address the following research questions in this paper:

RQ1: Do RBA-instrumented online services also use RBAR mechanisms?

RQ2: What RBAR challenges are used in practice?

RQ3: Are different RBAR challenges required when setting up MFA?

Contributions. This paper presents the first scientific insight into using RBAR in practice. We performed an exploratory analysis of RBAR behavior at Google and a systematic experiment on four other popular online services. We verified

⁴ To the best of our knowledge, there is no standard term for it yet.

RBAR at three of the five services. The analysis also included the influence of MFA configurations and different (virtual) locations. The main contributions achieved from these activities are the following:

- Identification of RBAR at popular online services
- A maturity model for different RBAR mechanisms

The remainder of this paper is structured as follows. Section 2 provides an overview of related work. In Section 3, we describe details behind how RBAR works. Section 4 explains the methodology of our experiments. The findings of the two experiments are described in Sections 5 and 6, respectively. Our overall results are discussed in Section 7. Section 8 summarizes our work and suggests possible future work.

2 Related Work

Most of the previous work on account recovery considered it a static mechanism. For instance, a lot of research focused on different additional authentication challenges for recovery that can be solved easily by legitimate users but not by potential attackers. Examples include cryptographic keys [9], delegated account recovery [20, 22], dynamic security questions [1, 19], and email address or phone number verification [26]. While these works do not address risk-based use cases, we argue that such methods would be beneficial in conjunction with a risk analysis of the user context.

Further research evaluated online services in the wild. Li et al. [23] studied the account recovery mechanisms of 239 popular online services in 2017 and 2019. They found that most of them implemented email address or mobile phone verification as a recovery mechanism. Amft et al. [6] conducted a large-scale study investigating which recovery methods are usually deployed in conjunction with MFA methods. They unveiled that website documentation usually does not correspond with the actual recovery procedure, showing the lack of transparency in account recovery. We confirm this as we analyzed the documentation of the services we tested for any references to RBAR, which in most cases were absent (see Section 7).

The only indication of risk-based recovery mechanisms we found in literature was mentioned by Bonneau et al. [7], where they noted that Google performed a “*risk analysis*” for account recovery. However, they did not further investigate how it works or what mechanisms are applied depending on the risk scenario.

Research on RBA is especially relevant for our work as it provides us with methods to analyze and develop risk-based systems. For example, Wiefeling et al. [38] studied RBA re-authentication mechanisms on five popular online services. They found that most online services used email verification to re-authenticate users. Gavazzi et al. [14] leaned on this work to identify that more than 75% of the 208 studied online services do not use any form of RBA. While the research in this field only addresses plain user authentication, our work extends

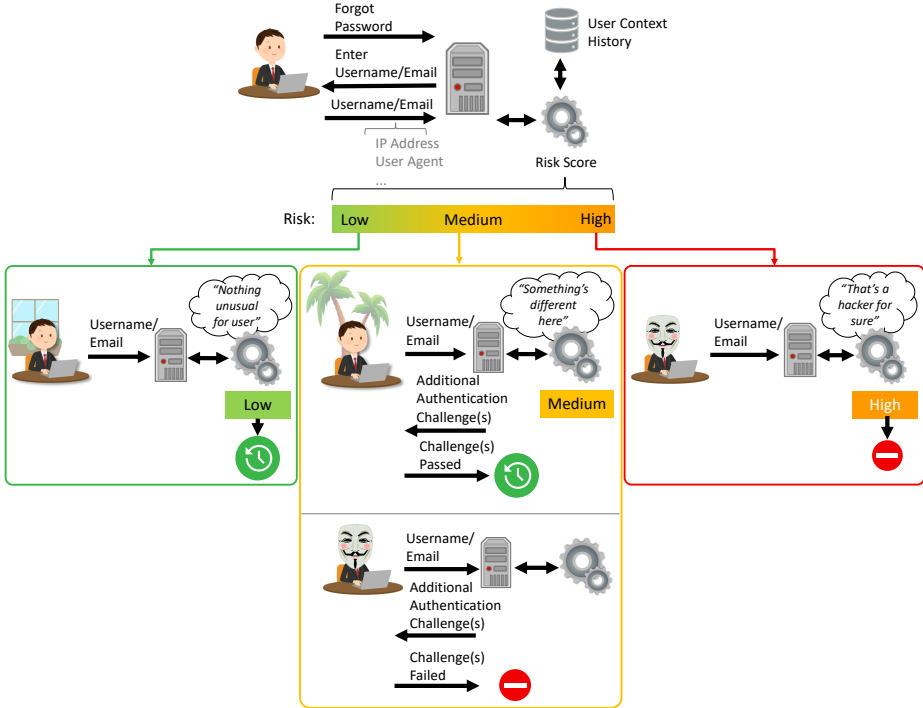


Fig. 1. Overview of the RBAR procedure (based on RBA illustration in [35])

it by showing that the methods used in RBA research can be equally applied in the context of account recovery. Consequently, we used the insights from prior work on RBA as a basis to study the use of RBAR on Google and other online services.

3 Risk-Based Account Recovery

Since there is no official description of RBAR yet, we describe its basic concept. Based on our observations on online services and previous knowledge in the related RBA field [36, 38], RBAR works as follows (see Fig. 1):

A user typically starts an account recovery process, e.g., by clicking “*forgot password*” at the online service’s login form. After that, the user is asked to enter the username or email used for the account to recover. While submitting this identifier, the user also submits additional feature data that is available in the current context to the online service, e.g., IP address or user agent string. Based on this information, RBAR compares these values with the user context history and calculates a risk score. The user context history contains feature values of past user actions, like previous legitimate logins that might have been validated by RBA mechanisms [37] before. The risk score is then classified into low,

medium, and high risk. Based on the risk, the online service performs different actions.

At a *low* risk, the feature values likely belong to the legitimate user, and the online service proceeds with the account recovery process (e.g., verify email address). A *medium* risk occurs if the user's feature values deviate from the expected values. The online service then introduces additional authentication challenges that require more user effort (e.g., solving a CAPTCHA or answering questions related to the account). After successfully solving these challenges, the online service proceeds with the account recovery process. A *high* risk means that the online service suspects that the user is likely targeted by a hacking attempt. The online service might block the account recovery process in these cases. However, to avoid locking out legitimate users trying to recover their accounts, this possibility has to be carefully selected by the online service.

4 Methodology

We investigated the research questions by conducting two experiments. Prior research has indicated that Google applies risk-based decision-making for account recovery [7], making it a suitable candidate for our first experiment. Therefore, we conducted an exploratory experiment on Google. We created test cases with different account setups, i.e., different authentication and recovery factor combinations. These were then tested with different user features to see how these could affect the recovery procedure. The study considered two RBA features, as suggested in Wiefeling et al. [38]: known/unknown browser and known/unknown IP address. A *known* browser is the one that was used before to sign in to Google, i.e., it has stored cookies from prior sessions. The *unknown* browser was tested using the browser's incognito mode to have a clean browser session without previously set cookies. The IP address feature was varied by using a VPN connection to be able to study the uncertain area of medium to high risk scores [38]. By comparing the recovery procedures of the different features for each test case, we identified the mechanisms used for RBAR. The test cases and the final results are given in Section 5.

For the second experiment, we developed an improved and more systematic approach. As the experiment required manual effort, we limited the number of tested services to the following services that are known to use RBA [14, 38]:

- LinkedIn (`linkedin.com`)
- Amazon (`amazon.com`)
- GOG (`gog.com`)
- Dropbox (`dropbox.com`)

The experiment was composed of three phases. First, we prepared user accounts for each service. Afterward, we checked whether any of the online services indicated RBAR behavior. Finally, since LinkedIn clearly turned out to implement RBAR, we analyzed if RBAR on LinkedIn is influenced by the MFA settings (as was the case with Google). More details on the steps and the results are presented in Section 6.

Table 1. Examples for Google account recovery without MFA enabled

Recovery factor	Phone signed in	Known browser	Known IP	Recovery procedure
None	○	●	●	Recovery not possible
None	●	●	●	1. Google prompt
None	●	○	○	1. Enter old password 2. Google prompt (two steps)
Email	○	●	●	1. Verify account email
Email	○	○	●	1. Enter old password 2. Verify account email

● = Feature present, ○ = Feature not present

5 Experiment 1: RBAR Use by Google

In the first experiment, we investigated previous assumptions [7] on whether Google used RBAR and identified features that might have an influence on the RBAR behavior. We describe the experiment and its results in the following.

5.1 Preparation

The exploratory experiment on Google was conducted between October 2021 and March 2022. We set up four Google user accounts that were created at intervals of several weeks to mitigate being detected as a researcher. Based on the visible feedback from the online service, we assume that we remained under the respective detection thresholds. In order to test the use of RBAR on Google, we defined the test cases based on the authentication and recovery factors offered in the Google account settings. At the time of the study, Google provided the following factors:

- **Main authentication:** password, sign in by phone
- **Secondary authentication:** Google prompt, phone call or text message, backup codes, security key, authenticator app
- **Recovery factors:** email, phone

The experiment on Google covered every possible single-factor authentication (SFA) account setup and eight MFA account setups. Each account setup was tested with all four RBA feature combinations. For each combination, all possible recovery options were explored.

5.2 Results

The study found that Google used RBAR for both SFA and MFA account setups. This became clear as using an unknown browser and/or an unknown IP

address increased the difficulty of recovering the account compared to using a known browser and IP address. This was indicated by requiring additional authentication factors, recovery options that were made unavailable, or an extra prompt like asking for the phone number of a registered phone.

Recovery Without MFA Enabled. Table 1 lists a few examples⁵ of the tests from studying SFA account recovery that clearly show the different recovery procedures based on RBA features. One can observe that in cases where an unknown browser was used for recovery, Google initially asked for an old password that the user could remember. This was not the case when using a known browser and a known IP address. The recovery procedure continued the same way, even if this step was skipped.

When a phone was signed in to the same Google account, this phone was prompted with a button showing “*Yes, it’s me*”. Users had to click this button to confirm the ownership of the account. This behavior changed when trying to recover the account from an unknown browser and an unknown IP address. In this case, Google also showed a two-digit number on the recovery web page and presented a dialogue with three number options on the phone. Users then had to select the correct number on the phone to proceed with the recovery.

Recovery With MFA Enabled. Table 2 shows some of the results that indicated obvious differences when trying to recover an account with a phone number configured for MFA. Note that in the given examples, we omitted the step of verifying access to the actual Google account email address to see what alternatives would be offered. When the recovery was performed from a known browser, it was sufficient to verify the phone that was set up for MFA by entering an OTP code that was sent to the phone via text message. Afterward, Google provided the user with an option to register and verify a new email address. A password reset email was sent to the newly registered email after 48 hours. In the meantime, the (legitimate) account owner got notifications about the ongoing recovery attempt. This allowed them to stop the procedure in case they did not request the recovery. However, this recovery option was not available when using an unknown browser. In that case, the user needed access to both the phone number and the email address registered on the actual Google account. This highlights how much RBAR features can impact the user’s chance of a successful recovery.

The last example in Table 2 shows a recovery procedure when using both an unknown browser and an unknown IP address. In this case, the user was first asked to enter the phone number used for MFA before actually verifying the ownership of this phone number.

Further Observations. Also, we observed that when failing a recovery, Google revealed some information on how its RBAR mechanism might work. The message displayed to the user on a failed recovery attempt suggested using a known device and Wi-Fi during recovery (see Fig. 2).

⁵ All results for the tests on Google are published on <https://github.com/AndreasTP/GoogleAccountRecovery>.

Table 2. Examples for Google account recovery with phone (text message) enabled for MFA

Recovery factor	Known browser	Known IP	Recovery procedure
None	●	● / ○	1. Verify MFA phone 2. Verify account email 3. Verify new email → Reset email after 48hrs
None	○	●	1. Verify MFA phone 2. Verify account email → Recovery not possible
None	○	○	1. Enter MFA phone number 2. Verify MFA phone 3. Verify account email → Recovery not possible

● = Feature present, ○ = Feature not present, ~~XXX~~ = Step omitted

However, during the study, we experienced that the recovery process could change from one day to another. This was true despite using the same account, having the same recovery options configured, and using the same browser and IP address. For instance, a recovery procedure that earlier gave access to the account after 48 hours through a password reset email ended in a failed recovery. An authentication factor that could previously be used to help recover an account was occasionally removed as a recovery option. This suggests that Google uses more RBAR features than the two tested in this study. Nonetheless, we confirm the assumption in prior work that Google implements a risk assessment in its recovery [7].

6 Experiment 2: RBAR Use by Other Services

The second experiment focused on online services that are known to use RBA [38] and investigated whether and how they also use some form of RBAR. We describe the experiment and its results below.

6.1 Preparation

For this experiment, we began by setting up user accounts for all four online services (see Section 4). Testing account recovery with personal accounts is not ideal since there is always the risk that accounts will be locked out or disabled entirely. However, RBA is oftentimes triggered only for legitimate accounts with a certain history of activity [38]. This makes sense from a technical perspective, as such algorithms need a certain amount of training data from the legitimate user to work correctly [37]. Therefore, we created four new test accounts for each of the services. These accounts were set up with the most basic settings, i.e., with

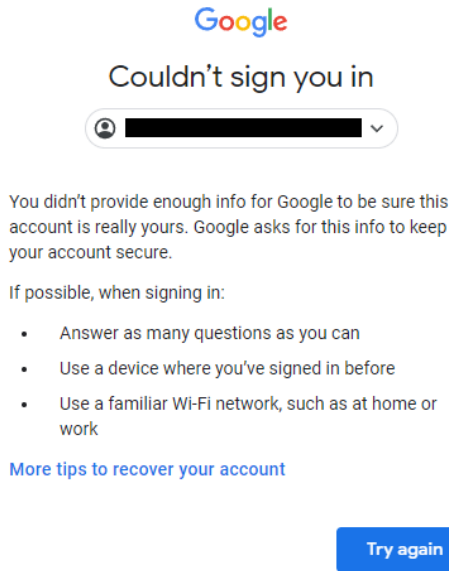


Fig. 2. Message shown when failing Google’s account recovery using an unknown browser and an unknown IP address. It reveals information that might give indications of their inner RBAR workings.

a password and one email address. To avoid bias, we made sure to create and use new email addresses on general-purpose email providers not linked to universities for each account. In addition, we were able to provide one old account for each service, some of which were either personal or created in previous studies.

A training was conducted in which the test accounts were logged in more than 20 times within a time period of about 1.5 months (December 2022–January 2023). We based the number of logins on Wiefeling et al.’s study [38]. Furthermore, it was ensured that the logins for each account were performed with a similar context, i.e., from the same browser and the same IP location. Also, logins from university IP addresses were avoided since experience has shown that online services might recognize these IP addresses and block accounts to prevent systematic analyses of their services. For reproducibility, we documented the context before each account login. We did this by recording all information from the IP address and HTTP header and the browser’s internal JavaScript functions, as in related work [36].

6.2 Identifying RBAR Usage

After training the test accounts, we analyzed whether the online services actually use RBAR mechanisms. As in related work by Wiefeling et al. [38] and Gavazzi et al. [14], this was tested by discovering differences in two distinct user contexts:

Table 3. Account recovery procedures for a normal and suspicious user context for the different test accounts of each online service

Online Service	Account	User context	
		Normal	Suspicious
Amazon	A1, A2, A4, A6*		EC
	A3, A1 [†]	CA → EC	CA → EC
	A5*	EC	<u>CA</u> → EC
Dropbox	D1 – D4, D5*		EL
GOG	G1 – G4, G5*	CA → EL	CA → EL
LinkedIn	L1 – L4, L5*	EC	<u>CA</u> → EC

EC = Email (Code), EL = Email (Link), CA = CAPTCHA,
** = Old account, † = Experiment repeated, XXX = Additional step*

normal and *suspicious*. This time, we considered a normal user to perform the account recovery from the same browser and IP location as in the training phase. In contrast, the suspicious user performs account recovery from a Tor browser. Web services can typically recognize Tor browser clients by the IP address of the exit nodes or by other browser features. Moreover, using a Tor browser is often considered suspicious [38]. We expected this to increase the likelihood of triggering risk-based mechanisms, if any, and compared to the first experiment on Google, where the Tor browser was not used. Note that we only considered differences that occurred after starting the recovery procedure for a specific account, e.g., after entering an email address. Any differences beforehand would not be relevant as it would mean that it is independent of the history of a user account.

Experimental Procedure. For this within-group experiment, account recovery was performed twice for each test account on different days at the end of January 2023, once with a normal user context and once with a suspicious user context, in varying orders, to avoid bias. This means we performed two account recoveries with all provided accounts. In the case of Amazon, we repeated the experiment with one of the new accounts and another old account due to inconsistent results, as described in more detail below.

Results. Table 3 summarizes the recovery procedures for each online service and account. Overall, the presentation of a CAPTCHA was the only noticeable difference that was found. The CAPTCHAs in the table are underlined in those cases where they appeared only in the suspicious user context. Note that Amazon uses its own AWS WAF CAPTCHA [5], while GOG uses the Google reCAPTCHA v2 [17] and LinkedIn appears to use a custom CAPTCHA implementation. Dropbox did not use any CAPTCHA within our experiments.

For **Amazon**, in three cases, only an OTP code sent via email was requested. Afterward, the password could be changed. For one of the new test accounts (A3), Amazon first requested a CAPTCHA before the email OTP code, but for both

normal and suspicious contexts. For the old account (A5), there was an actual difference as the CAPTCHA was only displayed in the suspicious context. Because of this inconsistent behavior, we did an additional test with A1, which this time required solving a CAPTCHA for both user contexts, similar to A3. Furthermore, we included a test with another personal account (A6) that was actively used to check if the behavior was related to the account age or activity. This time, no CAPTCHA had to be solved. Consequently, the risk assessment was more complex and could not be easily reproduced with our experimental setup.

Dropbox only requested the verification of the email address through a link before the password could be changed. This was the same for all user accounts, including the old one, and for both user contexts.

For **GOG**, a CAPTCHA had to be solved before verifying the email address through a link and finally changing the password. This was again equal for all accounts and both normal and suspicious user contexts.

LinkedIn was the only online service that consistently showed a different behavior depending on the context. For a normal user context, the email address had to be verified by an OTP code before the password could be changed. However, when performing recovery from a suspicious user context, a CAPTCHA had to be solved, sometimes multiple times.

In summary, Amazon and LinkedIn used RBAR, while Dropbox and GOG have not indicated any risk-based behavior during recovery. The only challenge that was shown depending on the user context was a CAPTCHA. The results for Amazon, however, were inconsistent for the different accounts. It was decided not to do a deeper analysis here, as the experimental setup clearly did not consider enough context parameters to simulate both a normal and a suspicious user context reliably. Yet, we conclude that Amazon must have used some form of RBAR. For LinkedIn, on the other hand, the RBAR behavior could clearly be reproduced with all accounts. Thus, we conducted a second experiment on LinkedIn using the newly created test accounts, as described in the subsequent section.

6.3 Analyzing the Influence of MFA Settings on Account Recovery on LinkedIn

In Section 5, we showed that Google implements RBAR by incorporating different authentication mechanisms that are set up as MFA factors in a user account. Since we could prove that LinkedIn also provides some form of RBAR, we conducted another experiment to determine whether LinkedIn used any other RBAR challenges beyond the CAPTCHA.

Experimental Procedure. For this experiment, we changed the authentication and recovery options in the LinkedIn test accounts. At the time of this experiment (January–February 2023), LinkedIn provided the following authentication and recovery methods:

- **Main authentication:** password

Table 4. Account recovery procedures for a normal and suspicious user context for the different LinkedIn account setups

#	Recovery		MFA		User context	
	Second Email	Text (SMS)	Auth. App	Text (SMS)	Normal	Suspicious
1	●	○	○	○	EC1 EC2	<u>CA</u> → EC1 EC2
2	○	●	○	○	EC1 P1	<u>CA</u> → EC1 P1
3	○	○	●	○	EC1 → AU	<u>CA</u> → EC1 → AU
4	○	○	○	●	EC1 → P2	<u>CA</u> → EC1 → P2
5	●	○	●	○	EC1 EC2 → AU	<u>CA</u> → EC1 EC2 → AU
6	●	●	○	●	EC1 EC2 → P2	<u>CA</u> → EC1 EC2 → P2
7	○	●	○	●	EC1 → P2	<u>CA</u> → EC1 → P2
8	○	●	●	○	EC1 P1 → AU	<u>CA</u> → EC1 P1 → AU

● = Feature present, ○ = Feature not present, EC1 = Primary Email (Code), EC2 = Secondary Email (Code), P1: Recovery Phone (SMS Code), P2 = MFA Phone (SMS Code), AU = Authenticator App, CA = CAPTCHA, | = Alternative XXX = Additional step

- **Secondary authentication:** phone (SMS), authenticator app
- **Recovery factors:** email address, phone (SMS)

We tested the effects of all possible combinations of these methods. In addition, LinkedIn also offered a non-digital recovery method requiring the user to submit a copy of a government-issued ID. As this would have revealed the experimenters’ identities, we did not include this method in the experiment. Similar to Google, the expected outcome for LinkedIn was that different authentication factors would be requested in a suspicious user context.

Results. Table 4 shows the results for the different tested account setups. Note that in setups 1, 2, 5, 6, and 8, there are two possibilities for receiving the verification code: as an alternative to the primary email address, the secondary email address or the phone number could be entered (indicated by the “|” symbol). LinkedIn allows configuring the same phone number as a second authentication factor and as a recovery method. In fact, when enabling the phone number for MFA, the same number is activated automatically for recovery by phone. However, in such cases, using the phone for account recovery does not make much sense as only a single factor (ownership of the SIM card) is required for resetting the password and logging in afterward, which contradicts the idea of *multi*-factor authentication. In these cases, i.e., setups 6 and 7, we only received an inaccurate error message (see Fig. 3). We filed a bug report for this to LinkedIn on February 24, 2023. However, the response from LinkedIn (one day later) indicated that it will not be fixed anytime soon unless it gets noticed by several other users.

The experiments show that the behavior when configuring further recovery or authentication methods is identical to the base setup. The only difference in the account recovery procedure for all setups was the initial CAPTCHA shown in the suspicious user context. Apart from that, the account recovery procedure always

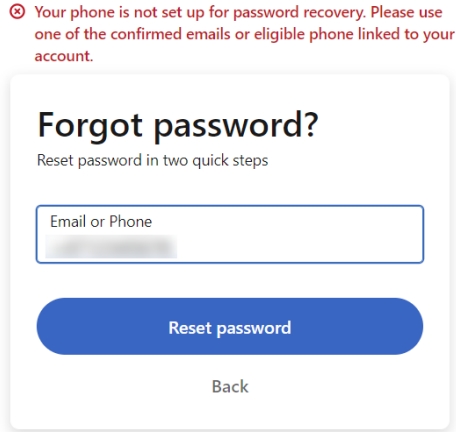


Fig. 3. Error message for phone recovery, when also Text Message MFA is activated

Table 5. Number of CAPTCHA iterations for different (pretended) locations for account recovery on LinkedIn

CAPTCHA iterations	Location of Tor exit
1	Sweden, Poland, United Kingdom, <i>Mexico</i>
2	United Kingdom, Germany
3	3× USA, <i>Czech Republic</i>
5	USA, Canada, <i>Netherlands</i>

started with the verification of the primary email address or phone number by an OTP code, followed by the verification of the MFA method if one was activated.

Variation of CAPTCHA Iterations. In addition to our main results, we observed that the number of iterations of the CAPTCHA on LinkedIn varied in different experiments between 1 and 5. When mapping the number of iterations to the pretended location (i.e., the location of the Tor exit node), an interesting correlation showed up (see Table 5). The normal usage location for all accounts was in Europe, and when the pretended location was also in Europe (just another country), 1 or 2 repetitions of the CAPTCHA were required. In cases where the suspicious location was on a different continent, 3 or 5 repetitions were needed. However, there were also cases (marked in italics) where this was not true. Nonetheless, it indicates that LinkedIn’s RBAR might give different suspicious risk classifications that are reflected in the number of CAPTCHA iterations. It also seems that the location is one important feature. Further experiments are needed to analyze to what extent other features are included.

7 Results and Discussion

In our exploratory study on Google and the follow-up experiment with other online services, we confirmed that several online services apply RBAR to a certain degree. In this section, we describe the results of the experiments with regard to the research questions. Furthermore, we summarize the results in a maturity model that we propose for RBAR implementations. Finally, we outline the limitations of our experiments and discuss further aspects of RBAR usage in practice.

7.1 Experiment Results

Within the scope of our experiments, we observed that Google implements RBAR in quite a sophisticated manner. It showed different authentication methods depending on the account setup and the user context. Dropbox and GOG did not apply any risk-based mechanisms during account recovery. Amazon actually indicated the use of RBAR, however, by assessing context information that was not considered by our two different user contexts. In some tests, a CAPTCHA had to be solved, while in others, it was not required. LinkedIn clearly behaved differently in a suspicious user context. When trying to recover an account from a Tor browser, LinkedIn showed a CAPTCHA challenge before entering an email verification code. In contrast to Google, however, the RBAR for LinkedIn did not involve MFA settings in a user account.

With regard to **RQ1**, we conclude that there are online services that use RBA, which also use RBAR—including Google, Amazon, and LinkedIn—but not all of them. To answer **RQ2**, the challenges we found on Google include pre-configured MFA methods (e.g., phone number) and questions requiring background knowledge (e.g., old passwords). On LinkedIn and Amazon, we only observed a CAPTCHA challenge in connection with RBAR. Concerning **RQ3**, we found that the MFA settings influenced the recovery procedure on Google only, while LinkedIn did not vary RBAR challenges depending on any configured MFA methods.

7.2 Maturity Model

Based on our results and inspired by [30], we propose a maturity model that ranks the different RBAR challenges by difficulty for an attacker (see Table 6). Due to the nature of RBAR, the model only considers the measures used in connection with a risk assessment. It describes the additional security gain in case the primary recovery factor (e.g. email address), if any, has already been compromised. Thus, no RBAR at all is considered the least mature as it does not involve any risk assessment and does not provide additional measures. Showing a CAPTCHA is ranked as level 1 as it can prevent automated attacks. Yet an attacker might bypass it or manually exploit account recovery. Background questions are ranked as level 2 as they require an attacker to gather knowledge of a victim. However, it also increases only the cost of the attack. MFA methods that

Table 6. Maturity model with maturity levels, mapping of RBAR challenges to the tested services and possible attacks against these challenges

Maturity	RBAR challenge	Identified on	Possible attacks
3	Pre-configured MFA	Google	Physical attack, malware [8]
2	Background knowledge	Google	OSINT, leaked passwords, phishing [1, 19]
1	CAPTCHA	LinkedIn, Amazon	Manual recovery, CAPTCHA bypass algorithm [21, 32]
0	None	Dropbox, GOG	n/a

are pre-configured in an account are considered the most mature as they require more sophisticated methods or even physical access for a successful attack.

The model can be used, e.g., to assess the security of an RBAR implementation. Online services can also use such a model for their RBAR implementations to enable certain challenges with a higher maturity ranking at higher risk scores. Note that the model is only one possible way to assess RBAR. It might be different if other types of RBAR challenges are used that were not discovered within our study.

7.3 Comparison with Official Documentation

To the best of our knowledge, the experiments showed for the first time that Amazon, LinkedIn, and Google use RBAR. To compare our findings with the public communications of the online services, we took their official documentation into consideration [4, 10, 15, 18, 25]. Interestingly, none of the RBAR-instrumented online services mentioned that they change the account recovery behavior based on contextual information collected during the recovery process [4, 18, 25]. Only Google hinted that users should possibly use a familiar device and location [18]. However, they did not mention why users should do this, i.e. because they use RBAR. Our results show that the account recovery mechanisms of these online services seem to do more to protect their users than what is officially communicated to them.

Trying to hide implemented security mechanisms from the user base has already been observed in the related case of RBA [16] and other research on account recovery [6]. We do not consider this a good practice, as it follows the anti-pattern of “*security by obscurity*”. Users also tend to get frustrated when they experience security barriers that were not communicated to them beforehand [34]. Beyond that, attackers are known to adapt to obscured security mechanisms [28, 33]. We assume that public RBAR research, to increase the body of knowledge, will increase the overall adoption of online services and enable a large user base to be protected with RBAR following the principle of “*good security now*” [13].

7.4 Ethics

We only tested account recoveries with accounts owned by the researchers, i.e., we did not try to exploit the recovery of other users' accounts. Also, since we conducted manual tests, we did not create high traffic on the online services that could have affected other users.

While it could be reasoned that our findings are helpful for attackers, we argue that they are more valuable to the public. As the gained knowledge helps researchers and online service providers to get an understanding of how RBAR works, this can support the development of more secure and usable account recovery mechanisms.

7.5 Limitations

Beyond Google, only four online services were analyzed in terms of RBAR. This was mainly due to the lack of any automatism for training user accounts and testing account recovery, therefore requiring manual effort to conduct our experiments. Nevertheless, as mentioned before, these services have been carefully selected as they are known to use RBA [38].

We could not find any RBAR mechanisms in Dropbox and GOG. Due to the nature of a black-box test, we do not know the implementation details of the tested online services. Thus, there is always uncertainty involved. Nevertheless, we are confident that the accounts were sufficiently trained—especially since we also tested older accounts—and tested with the highest risk possible [38].

7.6 RBAR

Attackers may abuse account recovery to circumvent authentication. Hence, the security of account recovery is as essential as the security of login authentication. Previous research showed that email addresses often become a single point of failure [23, 24]. RBAR might be an advantageous way to increase the difficulty of a successful account takeover by incorporating additional authentication methods, as with RBA. At the same time, it may reduce the burden on legitimate users and increase their chances of recovering an account.

The RBAR used by Google is quite different from LinkedIn. Google uses additional authentication methods, while LinkedIn just requires a suspicious user to solve an additional CAPTCHA. This CAPTCHA actually only reduces the risk of automated attacks by making it more costly for an attacker. In general, CAPTCHAs mainly increase friction for users [39]. It may be an improvement to use a risk score to decide if a CAPTCHA should be solved, compared to, e.g., GOG, where a CAPTCHA is shown to all users. However, the security gain is insignificant since researchers have already demonstrated attacks against Google's widely known reCAPTCHA [21, 32]. Moreover, this does not prevent targeted attacks. We argue that if a service already implements a risk assessment in its account recovery, it should even go further and include actual authentication methods. In the case of LinkedIn, it could, for instance, request the verification of another recovery email or phone if set up.

8 Conclusion

Account recovery mechanisms remain a relevant entry point for account takeover attacks [27, 29]. Online services should strengthen their account recovery with additional security mechanisms, like risk-based account recovery (RBAR), to protect their users.

In this paper, we investigated the use of RBAR in practice. We described the concept behind RBAR and conducted two experiments to learn about if and how online services use it. The results show that Google, Amazon and LinkedIn used RBAR. However, their implementations differed widely in suspicious contexts, from asking users for background knowledge or pre-configured MFA methods (Google) to showing a CAPTCHA challenge (Amazon and LinkedIn). Based on our results, we proposed a maturity model that researchers or service providers can use to assess the security of RBAR systems or guide in implementing RBAR.

Following this first systematic analysis of RBAR, future work can extend our proposed model with other RBAR challenges. Furthermore, it can be studied what features specifically trigger RBAR challenges. As there seems to be a tendency to include risk-based decision-making into account recovery, there should be a comparison of RBA and RBAR and how they can complement each other in authentication systems as a whole.

Acknowledgments Stephan Wiefeling did this research while working at H-BRS University of Applied Sciences.

References

1. Addas, A., Salehi-Abari, A., Thorpe, J.: Geographical Security Questions for Fallback Authentication. In: PST '19. IEEE (2019). <https://doi.org/10.1109/PST47121.2019.8949063>
2. Akamai: Credential Stuffing: Attacks and Economies. [state of the internet] / security 5(Special Media Edition) (2019), <https://web.archive.org/web/20210824114851/https://www.akamai.com/us/en/multimedia/documents/state-of-the-internet/soti-security-credential-stuffing-attacks-and-economies-report-2019.pdf>
3. Akamai: Loyalty for Sale – Retail and Hospitality Fraud. [state of the internet] / security 6(3) (2020), <https://web.archive.org/web/20201101013317/https://www.akamai.com/us/en/multimedia/documents/state-of-the-internet/soti-security-loyalty-for-sale-retail-and-hospitality-fraud-report-2020.pdf>
4. Amazon: Reset Your Password (2023), <https://web.archive.org/web/20210918230138/https://www.amazon.com/gp/help/customer/display.html?nodeId=GH3NM2YWEFEL2CQ4>
5. Amazon Web Services, Inc.: What is a CAPTCHA puzzle? (2023), <https://docs.aws.amazon.com/waf/latest/developerguide/waf-captcha-puzzle.html>
6. Amft, S., et al.: Lost and not found: An investigation of recovery methods for multi-factor authentication. In: arXiv:2306.09708 (2023)

7. Bonneau, J., Bursztein, E., Caron, I., Jackson, R., Williamson, M.: Secrets, Lies, and Account Recovery: Lessons from the Use of Personal Knowledge Questions at Google. In: WWW '15. ACM (2015). <https://doi.org/10.1145/2736277.2741691>
8. Campobasso, M., Allodi, L.: Impersonation-as-a-service: Characterizing the emerging criminal infrastructure for user impersonation at scale. In: CCS '20. ACM (2020). <https://doi.org/10.1145/3372297.3417892>
9. Conners, J.S., Zappala, D.: Let's Authenticate: Automated Cryptographic Authentication for the Web with Simple Account Recovery. In: WAY '19 (2019)
10. Dropbox: Change or reset your Dropbox password (2023), <https://web.archive.org/web/20230518113022/https://help.dropbox.com/security/password-reset>
11. Federal Bureau of Investigation: Internet Crime Report 2022 (Mar 2023), https://web.archive.org/web/20230311011752/https://www.ic3.gov/Media/PDF/AnnualReport/2022_IC3Report.pdf
12. Freeman, D., Jain, S., Dürmuth, M., Biggio, B., Giacinto, G.: Who Are You? A Statistical Approach to Measuring User Authenticity. In: NDSS '16. Internet Society (2016). <https://doi.org/10.14722/ndss.2016.23240>
13. Garfinkel, S.L.: Design Principles and Patterns for Computer Systems that are Simultaneously Secure and Usable. Ph.D. thesis, Massachusetts Institute of Technology (2005)
14. Gavazzi, A., et al.: A Study of Multi-Factor and Risk-Based Authentication Availability. In: USENIX Security '23. USENIX Association (2023)
15. GOG: How do I reset my password? (2023), <https://web.archive.org/web/20230317223608/https://support.gog.com/hc/en-us/articles/212185409-How-do-I-reset-my-password-?product=gog>
16. Golla, M.: I Had a Chat about RBA with @Google in April 2016. The Short Story: "RBA Is an Arms Race, and We Are Not Revealing Any Details That Could Potentially Help Attackers." (Apr 2019), <https://web.archive.org/web/20210812104239/https://twitter.com/m33x/status/1120979096547274752>
17. Google: reCAPTCHA v2 | Google Developers (2021), <https://developers.google.com/recaptcha/docs/display>
18. Google: Tips to complete account recovery steps (2023), <https://web.archive.org/web/20230422113749/https://support.google.com/accounts/answer/7299973>
19. Hang, A., De Luca, A., Hussmann, H.: I Know What You Did Last Week! Do You?: Dynamic Security Questions for Fallback Authentication on Smartphones. In: CHI '15. ACM (2015). <https://doi.org/10.1145/2702123.2702131>
20. Hill, B.: Moving Account Recovery beyond Email and the "Secret" Question. In: Enigma '17. USENIX Association (2017)
21. Hossen, M.I., et al.: An object detection based solver for google's image recaptcha v2. In: RAID '20. USENIX Association (2020)
22. Javed, A., Bletgen, D., Kohlar, F., Dürmuth, M., Schwenk, J.: Secure Fallback Authentication and the Trusted Friend Attack. In: ICDCSW '14. ACM (2014). <https://doi.org/10.1109/ICDCSW.2014.30>
23. Li, Y., Chen, Z., Wang, H., Sun, K., Jajodia, S.: Understanding Account Recovery in the Wild and its Security Implications. IEEE TDSC **19**(1) (2020). <https://doi.org/10.1109/TDSC.2020.2975789>
24. Li, Y., Wang, H., Sun, K.: Email as a master key: Analyzing account recovery in the wild. In: INFOCOM '18. IEEE (2018). <https://doi.org/10.1109/INFOCOM.2018.8486017>
25. LinkedIn: Password Reset Basics (2023), <https://web.archive.org/web/20221229120339/https://www.linkedin.com/help/linkedin/answer/a1382101>

26. Markert, P., Golla, M., Stobert, E., Dürmuth, M.: Work in Progress: A Comparative Long-Term Study of Fallback Authentication. In: USEC '19. Internet Society (2019). <https://doi.org/10.14722/usec.2019.23030>
27. Microsoft Detection and Response Team: DEV-0537 criminal actor targeting organizations for data exfiltration and destruction (2022), <https://www.microsoft.com/security/blog/dev-0537>
28. Milka, G.: Anatomy of Account Takeover. In: Enigma '18. USENIX Association (Jan 2018)
29. MITRE Corporation: CWE-640: Weak Password Recovery Mechanism for Forgotten Password (2021), <https://cwe.mitre.org/data/definitions/640.html>
30. Pöhn, D., Gruschka, N., Ziegler, L.: Multi-account dashboard for authentication dependency analysis. In: ARES '22. ACM (2022)
31. Quermann, N., Harbach, M., Dürmuth, M.: The State of User Authentication in the Wild. In: WAY '18 (Aug 2018), <https://wayworkshop.org/2018/papers/way2018-quermann.pdf>
32. Sukhani, K., Sawant, S., Maniar, S., Pawar, R.: Automating the bypass of image-based captcha and assessing security. In: ICCCNT '21. IEEE (2021). <https://doi.org/10.1109/ICCCNT51525.2021.9580020>
33. Thomas, K., et al.: Data Breaches, Phishing, or Malware?: Understanding the Risks of Stolen Credentials. In: CCS '17. ACM (2017). <https://doi.org/10.1145/3133956.3134067>
34. Wiefing, S., Dürmuth, M., Lo Iacono, L.: More Than Just Good Passwords? a Study on Usability and Security Perceptions of Risk-based Authentication. In: ACSAC '20. ACM (2020). <https://doi.org/10.1145/3427228.3427243>
35. Wiefing, S., Dürmuth, M., Lo Iacono, L.: Verify It's You: How Users Perceive Risk-based Authentication. *IEEE Security & Privacy* **19**(6) (2021). <https://doi.org/10.1109/MSEC.2021.3077954>
36. Wiefing, S., Dürmuth, M., Lo Iacono, L.: What's in Score for Website Users: A Data-Driven Long-Term Study on Risk-Based Authentication Characteristics. In: FC '21. Springer (2021). https://doi.org/10.1007/978-3-662-64331-0_19
37. Wiefing, S., Jørgensen, P.R., Thunem, S., Lo Iacono, L.: Pump Up Password Security! Evaluating and Enhancing Risk-Based Authentication on a Real-World Large-Scale Online Service. *ACM TOPS* **26**(1) (2023). <https://doi.org/10.1145/3546069>
38. Wiefing, S., Lo Iacono, L., Dürmuth, M.: Is This Really You? An Empirical Study on Risk-Based Authentication Applied in the Wild. In: IFIP SEC '19. Springer (2019). https://doi.org/10.1007/978-3-030-22312-0_10
39. Yan, J., El Ahmad, A.S.: Usability of captchas or usability issues in captcha design. In: Proceedings of the 4th symposium on Usable privacy and security. pp. 44–52 (2008)