A survey of user authentication systems and their revocability

Austin Kim
Computer Science Department
California State University – Fullerton
Fullerton, United States
dkim286@csu.fullerton.edu

***Abstract*—User authentication is a critical part of cloud computing. It is often the first step taken before conducting sensitive digital activities, such as unlocking one's personal device or accessing one's online service accounts. For this reason, authentication has been the go-to target for both researchers and malicious actors – the former seeking ways to secure the digital world, the latter seeking ways to do digital harm. Of interest to the security professionals is the former – the vast body of research and review of authentication mechanisms, both current and new. This paper aims to provide a survey of those studies by providing a brief overview of current authentication mechanisms as well as an analysis on how much user privacy they trade away for the sake of security. To achieve the latter, this paper introduces *Revocability*, a metric on how easy it is to invalidate breached secrets. Revocability is then used to assess the privacy impact of various authentication schemes and their potential harm in case of an inevitable data breach.**

***Index Terms*—Authentication, password, biometrics, 2FA, privacy**

I. Introduction

In the current era, it is virtually impossible to conduct daily routines and obligations without interfacing with digital devices and relying on online services. For example, roughly 7.3 billion business-to-business (B2B) transactions took place through the Automated Clearing House (ACH) network in the second quarter of 2021 alone, totaling up to $18.4 trillion for that quarter [1].

This puts a tremendous amount of burden on authentication systems: not only do they have to provide a reasonably secure way to transmit and receive secrets, they must ensure that whoever they're authenticating is in fact the legitimate user and not an adversary. Failure to do so would harm the user in multiple ways:

1) *Financial harm* – authenticating an adversary would subject the legitimate user to a serious financial harm [2]. Nearly all financial institutions today offer online banking via mobile phone apps or browser-based web apps. Breach of a mobile device's authentication system or that of a banking entity would leave the user's financial future at the mercy of an adversary.
2) *Harm to privacy* – smartphones carry an inordinate amount of personal and private data [3]. An unauthorized user (worse, an adversary) could reconstruct the footprints of the affected user for malicious purposes.
3) *Unexpected and unrelated harm* – most users reuse the same password across multiple accounts [4], thus potentially propagating the harm further to different and unrelated service accounts owned by the affected user.

Traditionally, research efforts into authentication mechanisms focused on the security and usability aspects, but they rarely delved into the privacy aspect of said mechanisms. Thankfully, this has been changing in recent years in the light of increased interest in safeguarding user data against actors who seek to exploit them for both profit and political reasons [5]. The concern is especially acute for biometric-based authentication systems, wherein the means of authentication (biometric data) is intrinsically tied to user identity. Breach of biometrics data would not only threaten the privacy of affected users, but it would also render those breached biometrics data forever unusable for authentication [2]. A significant enough breach of biometrics database would mean the cessation of one of the most convenient means of user authentication – a net disutility for society.

This paper aims to provide a survey of current authentication mechanisms. Perhaps more importantly, this paper focuses on the privacy implications of various authentication mechanisms. To achieve this, various authentication mechanisms are categorized into three categories discussed in [2]: authentication based on "what you know" (knowledge-factor), "what you are" (inherence-factor), and "what you own" (possession-factor). Furthermore, this paper proposes *revocability*, a measure of how easy it is for users to invalidate breached secrets. Then, various authentication mechanisms are analyzed using the revocability metric.

This paper is organized as follows:

- Section II covers the preliminaries – key concepts and terms related to authentication.
- Section III explains how authentication systems can be categorized.
- Section IV explains why privacy should matter for authentication mechanisms and introduces the revocability metric.
- Section V compares various authentication mechanisms based on their revocability.
- Section VI touches on outstanding issues and future research directions.

II. Preliminaries

A. Phases of Authentication

Broadly speaking, most authentication mechanisms operate in two separate phases: *enrollment phase* and *authentication phase*.

1) Enrollment phase: During the enrollment phase, an authorized user's secret (password, PIN), biometrics (fingerprint, iris scan), or token (smartcard, RFID badge) is registered to the authentication system. Then, depending on the authentication mechanism that the system utilizes, the entirety of the enrolled data is stored in some secure form (e.g. salted hash) or useful features are extracted and stored instead (biometric, continuous authentication) [6].

2) Authentication phase: During the authentication phase, this enrolled data is compared against the data provided by the user seeking to be authenticated. If the user's provided credential matches enrolled data to a degree of satisfaction, the user is granted access to the device or service.

B. Traditional vs. Continuous Authentication

There are two broad ways of instantiating authentication systems: *traditional authentication* and *continuous authentication*. This separation is unrelated to the factor-based *categorization* of authentication methods. In fact, authentication based on any factor can be used for either traditional or continuous authentication systems.

1) Traditional authentication: In traditional authentication, the user is authenticated only once in the beginning of the session. It requires the user to actively prove their authenticity to the system before they are allowed to access the device, service, or a physical space. Most traditional authentication systems utilize knowledge-factor-based authentication modes such as passwords.

2) Continuous authentication: In contrast, continuous authentication (CA) is done continuously and passively during a session using various forms of biometrics or context-aware authentication modes [6]. It monitors the user's behavior and the user's person throughout the session without any active authenticating effort from the user. Continuous authentication ensures that the authorized user is in control of the session, and revokes access if the authenticity of the user cannot be validated. Most continuous authentication systems utilize inherence-factor-based authentication modes such as browsing history or behavioral biometrics.

It is rare for systems to have CA as their sole authenticating mechanism. Rather, it's more often used in conjunction with traditional authentication mechanisms [6] to strengthen the security of the authenticating system as a whole.

C. Authentication Systems vs. Mechanisms

In some literature, *systems* and *mechanisms* are used interchangeably to describe authentication. This paper seeks to separate them into two distinct, dissimilar terms.

1) Authentication systems: This paper uses "systems" to refer to the instantiation of authentication mechanisms. For instance, an authentication system may employ password-based authentication mechanisms to verify users.

2) Authentication mechanisms: This paper uses "mechanisms" to refer to the ways through which users are authenticated. In some literature, this is referred to as an *authentication scheme*. This paper may use *mechanisms* and *scheme* interchangeably. For example, password-based authentication scheme would fall under the knowledge-factor-based category of authentication mechanisms.

III. Categorization of Authentication Systems

Categorizing authentication systems is not new. For example, the authors of [2] categorized authentication systems based on underlying factors employed in authentication: *knowledge* ("what you know"), *inherence* ("what you are"), and *possession* ("what you have"). Similarly, the authors of [6] categorized authentication systems based on a slightly different set of factors: *knowledge-based*, such as personal identification number (PIN) or password; *possession-based*, such as devices and smartcards; *physiological-based*, such as fingerprint and iris biometrics; *behavioral-based*, such as keystroke or touch patterns; *context-aware* such as physical location and browsing history. On the other hand, the authors of [7] chose to categorize authentication systems based on how they're interacted with and types of infrastructure needed to implement the authentication system: *password-based*, *biometric-based*, and *distributed* authentication systems.

This paper opts for the *factors*-based categorization scheme as described by [2].

This section is organized into subsections as follows:

- Section III-A briefly covers what *knowledge-factor* authentication mechanism is.
- Section III-B briefly covers what *inherence-factor* authentication mechanism is.
- Section III-C briefly covers what *possession-factor* authentication mechanism is.

A. Knowledge-factor

At its heart, *knowledge-factor* authentication mechanism challenges the user for something they know (e.g. password), who in turn answers the challenge question by proving their knowledge – by reproducing the piece of knowledge (e.g. type in the password) [2]. This category of authentication mechanism hinges on the simple assumption that only the valid user would possess the pre-enrolled knowledge. In an ideal scenario, no one else but the valid user would know the secret and be able to reproduce it for authentication.

Passwords and PIN numbers are perhaps the most famous authentication mechanisms belonging to the knowledge factor. Despite its age, there are new forms of knowledge-factor mechanisms being actively developed today, each utilizing recently proven or emerging technologies. Some of these new developments are outlined below. A list of other developments are in Table I.

TABLE I
KNOWLEDGE-FACTOR-BASED AUTHENTICATION MECHANISMS.

Authentication mechanism	Ref.
Spatial memory using map locations	[8]
GeoPass: map location-based password	[10]
IoT auth. questions based on user's past experiences	[11]
Questions based on user's smartphone geolocation data	[12]
HuMan: questions based on user's smartphone activity	[13]
Using episodic memory for user verification	[14]

TABLE II
INHERENCE-FACTOR-BASED AUTHENTICATION MECHANISMS.

Authentication mechanism	Ref.
Facial recognition (Apple *Face ID*)	[15]
Iris recognition (Samsung Galaxy S8)	[16]
Mobile phone auth. based on user's physical interaction with device	[17]
Smart-glass auth. based on user's head movement	[18]
Palmprint recognition	[19]

1) Spatial memory using map locations: In a study done by Hang et al [8], a location-based fallback authentication scheme was proposed. In it, the authors described a scheme where the user's spatial memory could be used as an authentication mechanism.

During the enrollment phase, the user chose some predefined locations to be used for authentication mechanism. Then, during the authentication phase, the user was asked to select the corresponding location on the map. Any location within 30 meters of the enrolled location was considered correct enough.

The study found that this scheme was reasonably strong against close adversaries, such as friends and partners. The downside of this mechanism was its high authentication time – up to 232 seconds in some extreme cases.

2) Graphical passwords: Suo et al [9] proposed a graphical password system that is resistant to shoulder-surfing and provides much better usability than traditional password-based authentication methods. In it, the authors presented an authentication mechanism that enrolled the user by allowing them to pick from a series of image thumbnails. Since images are more memorable than text phrases, this mechanism allowed image "passwords" to be longer, increasing its resiliency against dictionary attacks.

Despite the improved usability and longer average password length, it was found that this mechanism had some disadvantages. Mainly, it was found to be time-consuming to enroll new users and authenticate existing users.

B. Inherence-factor

The inherence-factor-based authentication mechanisms rely on either physical or behavioral aspects of the user to validate them [2]. Consequently, this factor is almost entirely the realm of biometrics-based authentication.

Physical inherence is about the physical characteristics of the user. It includes biometrics such as fingerprints, iris pattern, facial data, among others.

Behavioral inherence is about identifying patterns that are unique to an individual. It may include biometrics such as gait, voice, typing style, among others. Behavioral inherence has seen an uptick of interest with the proliferation of smartphones in the general population.

A selection of recent developments in inherence-factor-based authentication is described below. The rest are listed in Table II.

1) Facial recognition: Facial recognition-based authentication is already present in some mobile phones, including iPhones [15]. The iPhone variant of facial recognition, called *Face ID*, scans the user's face during the enrollment phase using its camera module. Then, during the login attempt, it takes an image of the user's face to make comparisons against the enrolled data.

Apple claims that Face ID is capable of making slight changes to the enrolled data to match the user's changing facial features as they age. It is also claimed to be nuanced enough to verify the user through glasses, change in clothes, headwear, and jewelry.

2) Iris recognition: Iris recognition-based authentication is also present in some mobile phones, including Samsung Galaxy S8 [16]. The Galaxy S8 variant of Iris recognition uses its camera and LED module to scan the user's iris for enrollment and authentication.

3) Mobile phone authentication based on user's digital activity: Fridman et al [17] proposed an authentication scheme that utilizes the various sensors on a smartphone to record the user's behavior and use that data for user verification. The scheme would collect a number of behavioral data from the user including typing behavior (misspelling, punctuation, etc), app usage, web visits (which URLs, how many times, when), and user's location info (GPS, WiFi). The data were collected over the course of 30 days from 200 subjects, and subsequently used for experimenting with this authentication scheme.

C. Possession-factor

The possession-factor-based authentication mechanisms rely on the authorized user having possession of a device or token [7]. Authentication schemes under this categorization require the user to first enroll a special device or token to the system, then to prove their ownership of the device or token during authentication phase.

The most common form of possession-factor-based authentication mechanisms include RFID cards and SMS one-time passwords (OTPs), the latter being used to prove the ownership of the enrolled mobile device. Possession-factor has become more widespread with the proliferation of mobile devices, as service providers do not need to consider the additional costs associated with distributing dedicated token devices.

Possession-factor alone is not used as the sole authentication method, since it is not considered to be secure enough. Instead, it is often employed in multi-factor authentication mechanisms as the secondary factor. One prominent example of this usage is two-factor authentication (2FA).

TABLE III
POSSESSION-FACTOR-BASED AUTHENTICATION SCHEMES.

Authentication mechanism	Ref.
Online service auth. using pre-enrolled GPS locations	[20]
YubiKey: dedicated device for generating OTP	[21]
Sound-Proof: ambient sound as 2FA	[23]
Unlocking Macbooks with Apple Watch	[24]

A selection of recent advances in possession-factor-based authentication are listed below. Table III contains a more complete (but not comprehensive) list of these mechanisms.

1) Pre-enrolled mobile GPS locations: Zhang et al [20] proposed an online service authentication scheme that uses a set of pre-enrolled GPS locations to ensure that the authenticating user is indeed in possession of their mobile device. The scheme enrolls the user by registering their mobile device and allowing them them to select a number of locations on the map where they plan to access the service from. Then, during authentication phase, the enrolled phone's mobile GPS data is checked to see if the user's mobile phone is in one of the enrolled locations. The user is validated only if they have their enrolled phone with them and is present at the enrolled location.

2) YubiKey – dedicated OTP device: YubiKey [21] is a dedicated physical device for generating OTPs. During enrollment phase, the user registers their OTP token device, such as YubiKey, to the authentication system. Then, during authentication phase, the system queries the user to prove their token ownership by entering the OTP generated on their token device. These OTPs are synchronized with the authentication system and only the correct OTP would result in the user being validated.

A number of major online services such as GitHub [22] support YubiKey for the possession-factor in their two-factor authentication schemes, along with other OTP token devices.

D. Multi-factor authentication (MFA)

Multi-factor authentication (MFA) is not an authentication category of its own. It refers to an authentication mechanism that combines two or more authentication mechanisms mentioned in the previous subsections.

This merger of multiple factors serves to strengthen the authentication system by providing some resiliency against attacks – if one factor becomes compromised for whatever reason, the remaining factor(s) may still remain uncompromised and prevent unauthorized access.

The most common form of multi-factor authentication is the *two-factor authentication* (2FA), where a knowledge-factor authentication (usually password) is combined with a possession factor (mobile device, YubiKey) to provide multi-factor resiliency.

IV. PRIVACY AND REVOCABILITY OF PRIVATE INFORMATION

This section covers why user privacy matters, what kind of harm may be caused by breach of privacy, and introduce the concept of *revocability* of private information.

A. Privacy and the Bottom Line

User privacy is not just an issue of individual benefit. Breach of user privacy can be detrimental to large organizations in terms of loss of revenue, loss of consumer confidence, and regulatory fines.

For example, breach of user privacy in the health sector alone costs $41.3 billion dollars annually [25]. In another example, Target suffered an infamous data breach in 2013 resulting from malware installed in the point-of-sale (PoS) systems, which ultimately cost the company more than $100 million in damages [26]. Perhaps the most infamous example comes from the Equifax breach incident, where nearly half of Americans' private information (name, address, SSN) were breached and leaked online.

B. Privacy and Physical Harm

Breach of user privacy does not stop at financial loss for the victim. For example, novel methods exist for correlating individuals' health with their credit scores alone [27]. Coupled with data breaches similar to the Equifax incident, it is within the realm of possibility for malicious actors to make their phishing attempts more effective (e.g. mention low-interest medical loans in the title) or downright blackmail them for nefarious purposes. Breach of privacy may also result in physical harm for users who live in under oppressive regimes [28].

C. Revocability of Private Information

A number of works describe mathematical privacy metrics based on various factors in exhaustive detail [29] [30].

For the sake of simplicity, this paper ranks the privacy impact of each revocability level with the following assumptions:

- A thorough data breach is inevitable.
- The breach will expose users' data in unencrypted, unhashed form.

This pessimistic posture is in-line with recent events [26] [31] and frequency of such events [4].

To briefly compare and contrast various authentication schemes, this paper examines the privacy impact incurred from the breach of private information organized into three different *revocability* levels:

- Impact of *revocable* private information
- Impact of *semi-revocable* private information
- Impact of *irrevocable* private information

1) Privacy impact of revocable information: Authentication information is considered revocable if it meets both of the following criteria:

- The user can take actions to render the information no longer up to date.
- Said action is simple, easy, and can be performed at no significant cost to the user.

For example, the user may change their password after an online service suffers a security breach, rendering the leaked password information no longer valid. Such revocations are effective for a single breach, but the fact that some users reuse

the same password across many online services [4] cannot be discounted. It is entirely possible that the leaked password remains valid for another online service that the user interacts with. Therefore, this paper considers breach of revocable information to have anywhere between *mild* to *moderate* impact on user privacy, depending on the authentication scheme.

2) Privacy impact of semi-revocable information: Authentication information is considered semi-revocable if it meets both of the following criteria:

- The user can take actions to render it somewhat out of date, but no further.
- Said action will incur a significant, but not insurmountable cost to the user.

Some possession-factor schemes would be considered semi-revocable as physical tokens and devices are expensive to replace and may result in other service accounts tied to that device being breached as well.

For example, knowing the user's name and phone number is often enough to obtain a password reset link from cellular carriers [32]. Once the attacker has the ability to intercept SMS messages, it can be used as a pivoting point to request password reset links from other online services or intercept SMS-based OTPs. The only way for the user to regain control would be to obtain a new phone number on short notice and re-enroll the new number to all the 2FA-enabled services belonging to them, which may be costly or nearly impossible to do in a short span of time. Therefore, this paper considers breach of semi-revocable information to have anywhere between *moderate* to *severe* impact on user privacy, depending on the authentication scheme.

3) Privacy impact of irrevocable information: Authentication information is considered irrevocable if it meets both of the following criteria:

- It is infeasible or impossible for the user to take actions that invalidate the leaked information.
- Even if it's feasible, it is prohibitively expensive or life-altering for the user.

For example, any data breach involving fingerprint-based biometrics system would require the user to surgically alter their fingerprints every time such incidents occurred. On the same token, any data breach involving behavioral biometrics that can't easily be changed (e.g. gait, handwriting) would be near impossible to change per-incident. Therefore, this paper considers breach of irrevocable information to have a *severe* impact on user privacy in most cases.

V. Privacy Comparison of Authenticating Information

A. Knowledge-factor Information

1) Passwords: The vast majority of password- and PIN-based schemes allow users to change their passwords or PINs, or even force the change unilaterally in case of a security breach. Therefore, passwords and PINs are revocable information.

TABLE IV
Privacy impact of knowledge-factor information.

Information Used	Revocability	Privacy Impact
Passwords	Revocable	Moderate
User's past events	Irrevocable	Severe

However, this does not necessarily mean that password- and PIN-based authentication mechanisms have low impact on user privacy. It is common for users to pick passwords from the names of people they know (family, friends, coworkers), choose PINs based on their street number or zip code, or prefix passwords with numbers that represent their birth year [33]. This is not an issue exclusive to the English-speaking world [34]. Exposure of such passwords would render it trivial for adversaries to bypass the affected user's accounts for different online services where security questions frequently ask for the same information as their knowledge-factor password (e.g. "What is your mother's maiden name?"). For this reason, password-based schemes have a moderate impact on user privacy.

2) User's past events: In some proposed works, the user's life experience is used as the knowledge-factor password for authentication [14] [11]. While this method does not directly use the user's data to generate these questions (rather, the questions and factoids are "enrolled" before authentication takes place), a breach of those stored factoids would be detrimental for user privacy.

Other methods exist that utilize the user's location history [10] [12] and smartphone activity [13] for forming knowledge-factor challenge questions. When breached, these schemes are detrimental to user privacy as well.

For example, something seemingly innocuous as zip codes can contribute significantly towards unmasking users. In one study, it was shown that a small set of data points including zip code, age, and gender contains enough identifiable information to fully identify 87% of US citizens [35].

As of this writing, there are no known ways to modify events that took place in the past. Therefore, the user's past events should be considered irrevocable information, whether provided voluntarily through enrollment [14] or monitored automatically for continuous authentication [12]. For the reasons stated, any authentication scheme that utilizes the user's past events should be considered to have a severe impact on user privacy.

B. Inherence-factor Information

1) Behavioral information: Of the discussed authentication schemes, smartphone usage and interaction [17] and head movement authentication [18] utilize behavioral information for direct authentication rather than using them to generate knowledge-based authentication questions.

Smartphone usage and interaction is semi-revocable. The location data is something that cannot be invalidated easily, as it would require the user to change their behavior or move to a different location. User's online browsing habit may

TABLE V
PRIVACY IMPACT OF INHERENCE-FACTOR INFORMATION.

Information Used	Revocability	Privacy Impact
Behavioral information	Semi-revocable	Moderate to Severe
Physical biometrics	Irrevocable	Severe

TABLE VI
PRIVACY IMPACT OF POSSESSION-FACTOR INFORMATION.

Device Used	Revocability	Privacy Impact
One-time password device	Revocable	Mild
One-time password via SMS	Semi-revocable	Moderate to Severe
Environment-aware device	Semi-revocable to Irrevocable	Moderate to Severe

be a factor that changes naturally with time, which would gradually invalidate the breached information. It may also be changed (with some difficulty) by willfully changing the user's browsing habit.

Even in best of cases where users' browsing data is stored in anonymized form, it has been shown to be trivial to de-anonymize the browsing data and correlate them to real users with nearly 70% accuracy [36]. Even datasets with only a small fraction of group membership data (e.g. which individuals is this person associated with?) is enough to successfully de-anonymize users [37]. Therefore, authentication schemes that use behavioral information have moderate to severe impact on user privacy.

2) Physical biometrics: Physical biometrics such as Apple's *Face ID* [15], Samsung's iris recognition system [16], and the more traditional palmprint recognition scheme [19] are firmly in the irrevocable side of private information. It is highly impractical or virtually impossible to change one's facial features every time a data breach occurs, and same goes for iris and palmprints. Therefore, authentication schemes that utilize physical biometrics have a severe impact on user privacy.

C. Possession-factor Information

1) One-time password device: A sizable chunk of possession-factor devices are centered around sending synchronized one-time passwords [21]. Such devices are costly to replace. Often, users utilize a single OTP device as the possession factor for nearly all of their multiple 2FA-enabled services. Many devices are in fact designed with this in mind and support multiple protocols.

Dedicated OTP devices are revocable. In case of loss or theft of such devices, it is a matter of simply replacing the device with a new one. Furthermore, disclosure of the device's PIN can be remedied by simply resetting the device at the cost of having to re-enroll the device to the user's 2FA-enabled service accounts. Therefore, authentication schemes that utilize a dedicated OTP device have a mild impact on user privacy.

2) One-time password on mobile phone (SMS): On the other hand, OTP via SMS is semi-revocable, as replacing a compromised mobile phone is far costlier in time and money. The fact that breaching SMS-based OTP necessarily involves SMS redirection [32] poses privacy issues as well. As long as the adversary has control over the user's phone number, they can read all incoming messages even if they are totally unrelated to the 2FA process, or perform password resets on the user's online service accounts to directly access private information. Therefore, authentication mechanisms that utilize OTP via SMS have a moderate to severe impact on user privacy.

3) Environment-aware device: Environment-aware devices utilize data points collected from the immediate surroundings of the user, which may pose a privacy risk by revealing the user's frequented locations.

One example [23] uses the ambient sound around the user's phone and the authenticating machine to verify that the user is in possession of the mobile device and is in proximity of the authenticating machine. Another example discussed previously [20] uses pre-enrolled GPS locations for two-factor authentication. The Apple Watch-based laptop unlocking scheme [24] uses the proximity of another device to grant physical access to a laptop.

While the physical devices being used as the second factor are revocable, the user's location habits are not as easily revocable. In the case of sound-based two-factor authentication described in [23] and location-based two-factor authentication, [20] any data breach related to their authentication location data would be virtually irrevocable if the user is mobility-challenged or lacks the means to authenticate from another location. On the other hand, a proximity-based second-factor such as Apple Watch is more easily revoked. Therefore, authentication mechanisms that utilize environment-aware devices have a moderate to severe impact on user privacy.

VI. ISSUES AND FUTURE DIRECTIONS

Research into privacy issues with regards to authentication systems is ongoing. However, much of the focus both in research and the industry is still on the security aspect of authentication systems. While it is laudable that large organizations are taking active steps to make their authentication systems more robust by phasing out password-only systems, [38] the issue of privacy must not be left to languish by the wayside. The same kind of design mistakes made around email metadata privacy should not be repeated again. [39] [40]

1) Biometrics and privacy: Biometrics-based authentication systems, whether it be inherence-based or behavioral-based, pose an acute threat to privacy. Loss or theft of biometrics data is catastrophic, as it cannot be revoked like passwords can. It'd require the user to quite literally modify parts of their physical and behavioral selves.

However, this does not mean that they're somehow inadequate or that they should be abandoned. Biometrics-based authentication systems are unmatched in usability and user acceptance. They're also backed by many years of research aimed at improving their accuracy and reducing false-positives. Instead, more research effort needs to be directed

TABLE VII
PRIVACY IMPACT COMPARISON OF ALL AUTHENTICATION FACTORS.

Factor	Information/Device Used	Revocability	Privacy Impact
Knowledge	Passwords	Revocable	Moderate
	User's past events	Irrevocable	Severe
Inherence	Behavioral information	Semi-revocable	Moderate to Severe
	Physical biometrics	Irrevocable	Severe
Possession	One-time password device	Revocable	Mild
	One-time password via SMS	Semi-revocable	Moderate to Severe
	Environment-aware device	Semi-revocable to Irrevocable	Moderate to Severe

towards preserving the users' privacy of these systems – specifically, how to retain their usability and accuracy while reducing their impact on privacy in case of data breaches.

2) Digital activity and privacy: Authentication systems that utilize the user's past activities – a form of inherence-based authentication – also pose an acute threat to privacy. Of the works reviewed in this paper, a worrying portion of them either glossed over the concern of user privacy or did not take active steps to consider user privacy in their authentication systems' architecture. The detailed extent to which these systems require the user to submit their day-to-day data is troubling. More research effort needs to be directed towards protecting the users' data in these authentication schemes.

VII. CONCLUSION

This paper delved into the categorization of authentication systems, mainly using their *factors*. Knowledge-factor authentication systems use the shared secret knowledge between the service (or device) and the user. Inherence-factor authentication systems use either physical or behavioral aspects of the user to authenticate them. Possession-factor authentication systems rely on the user to be in possession of the enrolled device, which is used either directly (RFID cards) or used to distribute one-time passwords (SMS-based OTP).

This paper also delved into user privacy and why it matters – how it can cause monetary harm to the user via loss of control over financial services, and also the potential for physical harm.

This paper also introduced the concept of *revocability* and how different levels of revocability may apply to different authentication schemes. It also provided a brief review of how each scheme rates in revocability and what kind of impact they may have in users' privacy.

Privacy and security in the authentication process are not mutually exclusive. It is entirely possible for an authentication mechanism to use something as intrinsically identifiable as inherence-factor factors to perform authentication and preserve users' privacy. As research continues into discovering new authentication mechanisms or shoring up the flaws of existing ones, there should be a renewed focus on preserving the privacy of users.

It is no longer a matter of opting into various online services for the sake of convenience. In this increasingly digitized world, interacting with technologies – and authenticating to them – is a necessity. This obligation should not come at the cost of individuals' privacy.

REFERENCES

[1] "ACH network volume and value statistics." [Online]. Available: https://web.archive.org/web/20210923001915/https://www.nacha.org/content/ach-network-volume-and-value-statistics

[2] S. W. Shah and S. S. Kanhere, "Recent trends in user authentication – a survey," *IEEE Access*, vol. 7, p. 112505–112519, 2019.

[3] K. Kirkpatrick, "Who has access to your smartphone data?" *Communications of the ACM*, vol. 63, no. 10, p. 15–17, Sep 2020.

[4] N. Lord, "Uncovering password habits: Are users' password security habits improving? (infographic)," Sep 2020. [Online]. Available: https://web.archive.org/web/20211027155119/https://digitalguardian.com/blog/uncovering-password-habits-are-users-password-security-habits-improving-infographic

[5] K. v. d. Schyff, S. Flowerday, and S. Furnell, "Duplicitous social media and data surveillance: An evaluation of privacy risk," *Computers Security*, vol. 94, p. 101822, Jul 2020.

[6] A. F. Baig and S. Eskeland, "Security, privacy, and usability in continuous authentication: A survey," *Sensors*, vol. 21, no. 17, p. 5967, Sep 2021.

[7] X. Wang, Z. Yan, R. Zhang, and P. Zhang, "Attacks and defenses in user authentication systems: A survey," *Journal of Network and Computer Applications*, vol. 188, p. 103080, Aug 2021.

[8] A. Hang, A. D. Luca, M. Smith, M. Richter, and H. Hussmann, "Where have you been? using location-based security questions for fallback authentication," in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*. Ottawa: USENIX Association, Jul. 2015, pp. 169–183. [Online]. Available: https://www.usenix.org/conference/soups2015/proceedings/presentation/hang

[9] X. Suo, Y. Zhu, and G. Owen, "Graphical passwords: a survey," in *21st Annual Computer Security Applications Conference (ACSAC'05)*, 2005, pp. 10 pp.–472.

[10] J. Thorpe, B. MacRae, and A. Salehi-Abari, "Usability and security evaluation of geopass: A geographic location-password scheme," in *Proceedings of the Ninth Symposium on Usable Privacy and Security*, ser. SOUPS '13. New York, NY, USA: Association for Computing Machinery, 2013. [Online]. Available: https://doi.org/10.1145/2501604.2501618

[11] N. Shone, C. Dobbins, W. Hurst, and Q. Shi, "Digital memories based mobile user authentication for iot," in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, 2015, pp. 1796–1802.

[12] Y. Albayram, M. Khan, A. Bamis, S. Kentros, N. Nguyen, and R. Jiang, "A location-based authentication system leveraging smartphones," vol. 1, 07 2014, pp. 83–88.

[13] P. Gupta, T. K. Wee, N. Ramasubbu, D. Lo, D. Gao, and R. K. Balan, "Human: Creating memorable fingerprints of mobile users," in *2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, 2012, pp. 479–482.

[14] S. S. Woo, R. Artstein, E. Kaiser, X. Le, and J. Mirkovic, "Using episodic memory for user authentication," *ACM Transactions on Privacy and Security*, vol. 22, no. 2, p. 1–34, Apr 2019.

[15] G. Fleishman, "Face id on the iphone x: Everything you need to know about apple's facial recognition," Dec 2017. [Online]. Available: https://www.macworld.com/article/230490/face-id-iphone-x-faq.html

[16] "Iris recognition on galaxy s8," Jan 2018. [Online]. Available: https://www.samsung.com/au/iris/

[17] L. Fridman, S. Weber, R. Greenstadt, and M. Kam, "Active authentication on mobile devices via stylometry, application usage, web browsing,

and gps location," *IEEE Systems Journal*, vol. 11, no. 2, pp. 513–521, 2017.

[18] S. Li, A. Ashok, Y. Zhang, C. Xu, J. Lindqvist, and M. Gruteser, "Whose move is it anyway? authenticating smart wearable devices using unique head movement patterns," in *2016 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2016, pp. 1–9.

[19] G. Jaswal, A. Kaul, and R. Nath, "Multiple feature fusion for unconstrained palm print authentication," *Computers Electrical Engineering*, vol. 72, pp. 53–78, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790617321419

[20] F. Zhang, A. Kondoro, and S. Muftic, "Location-based authentication and authorization using smart phones," in *Proceedings of the 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications*, ser. TRUSTCOM '12, 2012, p. 1285–1292. [Online]. Available: https://doi.org/10.1109/TrustCom.2012.198

[21] "Meet the YubiKey," Jun 2021. [Online]. Available: https://www.yubico.com/why-yubico/

[22] "Configuring two-factor authentication." [Online]. Available: https://docs.github.com/en/authentication/securing-your-account-with-two-factor-authentication-2fa/configuring-two-factor-authentication#configuring-two-factor-authentication-using-fido-u2f

[23] N. Karapanos, C. Marforio, C. Soriente, and S. Capkun, "Sound-proof: Usable two-factor authentication based on ambient sound," in *24th USENIX Security Symposium (USENIX Security 15)*. Washington, D.C.: USENIX Association, Aug. 2015, pp. 483–498. [Online]. Available: https://www.usenix.org/conference/usenixsecurity15/technical-sessions/presentation/karapanos

[24] "Unlock your mac with your apple watch," Nov 2020. [Online]. Available: https://support.apple.com/en-us/HT206995

[25] L. Guo, C. Zhang, J. Sun, and Y. Fang, "A privacy-preserving attribute-based authentication system for mobile health networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 9, p. 1927–1941, Sep 2014.

[26] N. Manworren, J. Letwat, and O. Daily, "Why you should care about the target data breach," *Business Horizons*, vol. 59, no. 3, p. 257–266, May 2016.

[27] E. A. Knapp and L. T. Dean, "Consumer credit scores as a novel tool for identifying health in urban u.s. neighborhoods," *Annals of Epidemiology*, vol. 28, no. 10, p. 724–729, Oct 2018.

[28] I. C. Campbell, "The Taliban may have seized biometric data that can ID US allies in Afghanistan," Aug 2021. [Online]. Available: https://www.theverge.com/2021/8/18/22630686/biometric-data-afghanistan-taliban-hiide-civilians

[29] C. Diaz, "Anonymity metrics revisited," 01 2005.

[30] I. Wagner and D. Eckhoff, "Technical privacy metrics: A systematic survey," *ACM Computing Surveys*, vol. 51, no. 3, p. 1–38, Jul 2018.

[31] "Chinese hackers charged in equifax breach," Feb 2020. [Online]. Available: https://www.fbi.gov/news/stories/chinese-hackers-charged-in-equifax-breach-021020

[32] "Anatomy of a hack." [Online]. Available: https://www.theverge.com/a/anatomy-of-a-hack

[33] X. Yu and Q. Liao, "Understanding user passwords through password prefix and postfix (p3) graph analysis and visualization," *International Journal of Information Security*, vol. 18, no. 5, p. 647–663, Oct 2019.

[34] C. Shen, T. Yu, H. Xu, G. Yang, and X. Guan, "User practice in password security: An empirical study of real-life passwords in the wild," *Computers Security*, vol. 61, p. 130–141, Aug 2016.

[35] L. Sweeney, "k-anonymity: A model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 05, p. 557–570, Oct 2002.

[36] J. Su, A. Shukla, S. Goel, and A. Narayanan, "De-anonymizing web browsing data with social networks," in *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, Apr 2017, p. 1261–1269. [Online]. Available: https://dl.acm.org/doi/10.1145/3038912.3052714

[37] G. Wondracek, T. Holz, E. Kirda, and C. Kruegel, "A practical attack to de-anonymize social network users," in *2010 IEEE Symposium on Security and Privacy*. IEEE, 2010, p. 223–238. [Online]. Available: http://ieeexplore.ieee.org/document/5504716/

[38] A. Simons, "A breakthrough year for passwordless technology," Dec 2020. [Online]. Available: https://www.microsoft.com/security/blog/2020/12/17/a-breakthrough-year-for-passwordless-technology/

[39] C. Soghoian, "Surveillance and security lessons from the petraeus scandal," Nov 2012. [Online]. Available: https://www.aclu.org/blog/national-security/privacy-and-surveillance/surveillance-and-security-lessons-petraeus-scandal

[40] "Email privacy concerns," Jun 2016. [Online]. Available: https://www.findlaw.com/consumer/online-scams/email-privacy-concerns.html