Haasnoot *et al.*
EURASIP Journal on Information Security (2022) 2022:5
https://doi.org/10.1186/s13635-022-00131-y

EURASIP Journal on Information Security

RESEARCH Open Access

Presentation attack detection and biometric recognition in a challenge-response formalism

Erwin Haasnoot*, Luuk J. Spreeuwers and Raymond N. J. Veldhuis

Abstract

Presentation attack detection (PAD) is used to mitigate the dangers of the *weakest link problem* in biometric recognition, in which failure modes of one application affect the security of all other applications. Strong PAD methods are therefore a must, and we believe biometric challenge-response protocols (BCRP) form an underestimated part of this ecosystem. In this paper, we conceptualize what BCRPs are, and we propose a descriptive formalism and categorization for working with them. We validate the categorization against existing literature that we classified to be describing BCRPs. Lastly, we discuss how strong BCRPs provide advantages over PAD methods, specifically in the protection of individual applications and the protection of *other* applications from inadvertent leaks in BCRP applications. We note that research in BCRPs is fragmented, and our intent for the proposed formalism and categorization are to give focus and direction to research efforts into biometric challenge-response protocols.

Keywords: Biometric challenge-response, Interactive biometrics, Presentation attack detection, Biometric recognition

1 Introduction

Biometrics are being applied on a larger and larger scale, but centralized databases that fuel biometric applications are always at risk of leaking data. Cases that relatively recently hit the news are the Suprema Biostar 2 leak [1] (2.8 million records) and the Aadhaar leak [2] (potentially 1.1 billion records). We should expect that leaks at this scale will happen again in the future and that when they do, we might not even hear about it.

Distribution of the leaked biometric data is already trivial and effectively cost-free with current internet bandwidth availability. For example, the Suprema Biostar 2 leak (23 gigabytes of data) can be transferred over the internet with a typical consumer grade internet connection in a matter of minutes. These speeds are only expected to increase in the future.

Plus, with the coming-of-age of anonymity focused internet protocols, such as The Onion Router (TOR) [3], and cryptographic currencies such as Bitcoin [4] or biometric currencies such as Worldcoin [5], anonymous messaging, market places, file transfers, and payments are available to anyone with an internet connection. This allows for anonymous monetization of biometric data leaks, incentivizing the illegal acquisition of said biometric data.

With the distribution, acquisition, and monetization of biometric data for malicious purposes becoming increasingly common and effective, as well as the growth of biometric databases in number, size and detail, an increasing problem for all biometric applications is that they are inherently linked; they can only "tap" from a limited, shared, non-revocable resource pool. The number of, e.g., irises a person can genuinely present to a sensor is limited, and it is not possible to "revoke" an iris and change it for a new one in response to a data leak.

This results in a *weakest link problem* where, without counter measures, individual biometric applications can only offer security guarantees as strong as the weakest out of all linked applications. Malicious actors with increasingly easy access to all biometric data will

*Correspondence: e.haasnoot@utwente.nl

Data Science Group, Faculty of EEMCS, University of Twente, Enschede, The Netherlands

be able to present artifacts with biometric characteristics that match any user that they want to impersonate and will thus pose significant risk to applications using biometric recognition, even when those have followed good security practices.

Under the banner of presentation attack detection (PAD), a concerted effort is being made to mitigate the effects of biometric data leaks and make biometric recognition more robust against leak-related failure modes. PAD methods attempt to harden individual applications against the failure modes of other applications. For example, by testing for the existence of a pupil reflex [6] a print of an (illegitimately retrieved) iris scan could be detected. This type of PAD is specifically known as liveness detection, as defined in ISO-30107-1 [7]. We believe the right kind of challenge-response protocols provide an additional path forward in that effort.

Challenge-response protocols are protocols where the authenticity of, e.g., persons can be ascertained through a "question" and "answer" style interaction. In biometrics, they have mostly been used as a form of liveness detection and provide hardening of individual applications. Additionally, we believe the right kind of challenge-response protocols can be applied in biometric recognition and as such provide additional protections in the case of the aforementioned *weakest link problem*.

Biometric challenge-response protocols have potential, but contemporary research into biometric challenge-response protocols is unfortunately fragmented. The aim of this paper is therefore to introduce a set of concepts relevant to the application of challenge-response protocols in biometrics. Specifically, we introduce a descriptive formalism and a categorization intended to augment challenge-response protocol-related concepts of ISO-30107-1, to clear up what the right kind of challenge-response protocols are and to provide a direction for more focused research into biometric challenge-response protocols.

We first discuss presentation attack detection and biometric recognition verification questions and their similarities. Second, we discuss in more detail what makes challenge-response protocols tick and introduce a definition for what we call biometric challenge-response. Third, we introduce the formalism and categorization. Fourth, we provide a validation of the categorization by way of a survey of the current state of literature of biometric challenge-response protocols. Finally, we will discuss what the right kind of challenge-response protocols are to achieve the goal of protecting the application collective from the individual applications. This, we hope, will focus and give direction to research efforts into biometric challenge-response protocols.

2 Presentation attack detection and biometric recognition

ISO-30107-1 defines presentation attack detection as the automated determination of whether presentations to a biometric system have the goal of interfering with the correct operation of said system. Biometric recognition (BR) concerns itself with the automation of matching (or explicitly not matching) instances of recorded biometric characteristics. PAD and BR concern themselves with different, but complementary, parts of a biometric system. Similarities are found when describing modes of operation of both PAD and BR, as we shall see below.

2.1 Modes of operation

In PAD, there are roughly 2 modes of operation. One is to detect specific features of a biometric presentation that are known to be hard to replicate in certain attack scenarios. This can be the way light reflectance changes or how certain marks move through 3-D space [8], as a user moves their head. Another mode is to try to detect features of known attack scenarios. This can, e.g., be detecting the idiosyncratic light reflectance from 3-D face masks. In terms of ISO-301071, the former can be done through, e.g., liveness detection and video surveillance, whereas the latter can be anything from artifact detection, alteration detection to coercion detection, and obscuration detection.

Applications that make use of biometric recognition capture a sample from some biometric source and then process this sample in one of three comparison modes. Mode one is *1:1* comparison, where the biometric *sample* is compared to only one other claimed identity, which can be in the form of some previously stored *template*. This comparison mode is also known as *verification* and is used in biometric authentication to check whether a person is who they claim to be. The second mode is *1:n* closed-set comparison, where a *probe* is compared to a biometric *reference/samples* database, to see whether there is a match to one of said samples. This is the mode often employed in identification, to find out the identity of the person being presented to the system. The third mode is *1:n+1*, or open-set, comparison mode. This mode is used for deduplication, where one needs to find out if a person is in some database (and who it is), or if that person is unknown, so that no person is enrolled into the deduplicated database twice.

In both PAD and BR, we can say the general mode of operation is establishing authenticity. This is, respectively, the authenticity of the modality, the trait from which the biometric data is derived, in PAD and of the identity in BR. So, the first mode described for PAD maps neatly onto BR's verification mode. The identification mode of BR maps on to the attempts at specifically detecting what type of fake is

being used. Table 1 shows the relation between the types of biometrics and its modes of operation. For this section, the focus will be on the 1:1 comparison mode, as this is the mode amenable to challenge-response protocols.

2.2 Formalizing PAD and biometric recognition

Framing PAD and biometric recognition modes of operation in a unified representation like this has the major advantage that we can express the questions we try to answer in both modes of operation in the same way. The verification mode question in PAD can be stated as *Is the user a real person?*, whereas the verification mode question in biometric recognition can be stated *Is the user person X?*. We would put the unifying verification question as such:

`Is the user really person X?`

Answers to this question are contextual to the biometric systems (or models) that are used to answer them. If the biometric system only does recognition, without any PAD, then an answer to the verification question is a statement on the identity of the user, whereas an answer in a system that implements both recognition algorithms and PAD mechanisms is a statement on both the identity and the liveness of the user. Pushing the "meaning" of the verification answer down into the system/model allows us to express what a biometric system does in a concise formalism.

A source S presents their biometric characteristic, which results in data X, see Eq. 1. A right arrow "→" indicates a transformation or assignment into a new variable, and an equals sign "=" indicates equality.

$$S \rightarrow X \tag{1}$$

In this context, we consider X a structure from which information about both the Modality I_M, which is useful for PAD, and about the Identity I_I can be extracted, see Eq. 2. Furthermore, X may also contain "Other" information, e.g., on the emotional state of a user, but for BR or PAD this is not usually of interest.

$$X = \{I_M, I_I, I_{\text{Other}}\} \tag{2}$$

Table 1 Model of how biometric recognition and PAD share the same modes of operation and where a non-exhaustive list of different applications fall in this model

Mode	Recognition	PAD (modality)
1:1	Authentication	Liveness Detection
1:n	Identification	Artifact Detection
1:n+1	Forensics, Deduplication	

Handling data as a structure like X makes sense from a system perspective, as the system needs to guarantee the integrity and security of *all* of X, and not just parts.

To extract and validate the information in X, we require a structured way to store our knowledge of the problem domain, which we will call Model M. M will be generated from both, previously acquired, modality and identity information, see Eq. 3.

$$\{I_M, I_I\} \rightarrow M \tag{3}$$

This information can come from two types of sources, specifically, the biometric source S, which will likely be represented by a historical collection of Xs, and other non-specific sources, which could for example be the general description of the light reflectance properties of skin taken from an academic paper. Modality information can be derived from all sources, but S is the only one that can provide identity information.

The formalism is then brought together by the expression of the *verification* question as whether the output of function F, that transforms data X using a model M into a similarity score, is above some tunable threshold T. This results in a matching decision D. See Eq. 4.

$$F(X, M, T) \rightarrow D \tag{4}$$

If we constrain $I_M = \emptyset$ during generation of M, F will become a scoring function for recognition, meaning we can describe "naive" applications of biometric recognition under this limitation. On the other hand, if we constrain $I_I = \emptyset$, F will become a pure PAD scoring function, allowing us to describe applications that purely consist of PAD as well. These types of applications most prominently include the enrollment phases of biometric authentication, where the identity is newly created or determined through some other method, and we only need to be assured that the modality is genuine.

A descriptive formalism like the above is useful, because it allows us to discuss biometric systems holistically without having to deconstruct it into PAD or BR implementations. It is also amenable to straight-forward extensions, e.g., to formalize biometric challenge-response protocols as we will show further on.

2.3 How PAD protects the application

We can use the descriptive formalism to understand how PAD protects applications from other weak implementations. Suppose a user, a source S_1, makes use of three biometric applications. Application 1 is leaking data X_1 containing I_{I_1} of S_1. Application 2 is not leaking data, but has not implemented a PAD method, such that its model M_2 is derived from $\{I_M = \emptyset, I_{I_1}\}$. Application 3 is not leaking data and also implements a PAD method. Its

model M_3 is derived from $\{I_M \neq \emptyset, I_{I_1}\}$. Figure 1 shows a representation of the applications.

A malicious agent retrieves X_1 that is leaking from application 1. It then creates a new "fake" biometric source S_1^* from this data X_1. For example, X_1 can be a photo of an iris and the artifact a print of this photo. The source S_1^* is presented to a sensor and results in new data X_1^* that contains the same I_I as the original X_1, but it lacks all the general information related to the modality I_M.

Application 2, with model M_2, will not be able to distinguish X_1^* from a genuine X_1, because its model lacks information about what the modality should look like. It should be considered compromised due to application 1's leaks. Application 3 with M_3 is more complex and does include information about the modality, e.g., it will be able to test for the pupil reflex. Application 3 is therefore more difficult to compromise, due to the workings of PAD.

Unfortunately, no PAD method is perfect. A malicious agent could figure out a way of recreating the modality information in the "fake" source S_1^*, such that it would bypass the security measures taken by application 3. A photo could be printed on a material that has the same light reflectance properties of skin (although perhaps not the right 3D geometry) or is stretchy and able to deform and thus able to simulate a form of pupil dilation. The agent would be able to gather this information from general sources or even data leaked from PAD-implementing applications.

It would be useful to have PAD schemes that would allow for some general information to be used some of the time and for other general information to be actively harmful to the odds of successfully triggering false matches in the application. Introducing such a verification session-based component to the matching process is where challenge-response protocols find their application in biometrics.

3 Challenge-response protocols

Challenge-response protocols are applied in a wide range of fields in as many forms as there are fields but can be defined as the family of protocols that allow one party A to present a challenge (the "question") to which a party B must respond (the "answer"), such that the authenticity of party B's claim can be established by party A. More extensive challenge-response protocols might allow party B to challenge party A in a similar manner, such that mutual authenticity can be established. Examples of challenge-response protocols in fields outside of biometrics are digital signature algorithms and viva voce (oral) exams.

In digital signature algorithms, document-specific digests are processed into digital signatures so that the authenticity of said document can be tested. In viva voce exams, domain-specific questions are processed into relevant answers to test whether a student truly possesses some corpus of knowledge. Figure 2 shows an example of a viva voce interaction that happens when a student Stu wants to convince biometrician Bob he belongs to the exclusive club of biometricians.

3.1 Process asymmetry

Challenge-response protocols rely fundamentally on an asymmetry in the process that transforms the challenge into a response. It should be easy to verify that a certain response belongs to a specific process but very hard to

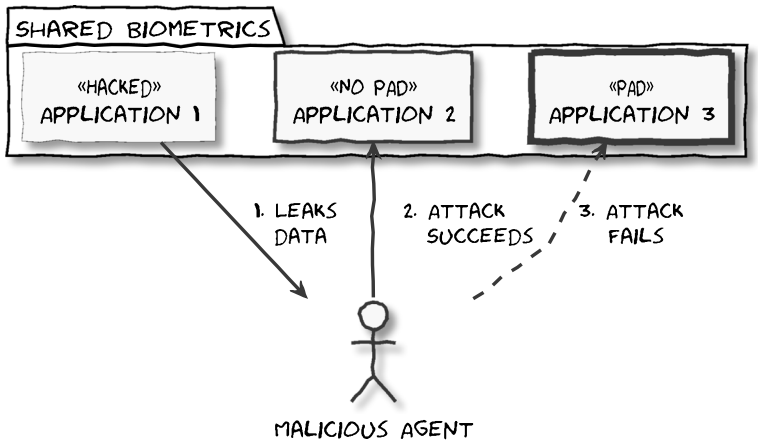

Fig. 1 A selection of three applications that all share the same biometric source. Application 1 is lacking in security and leaking data to a malicious agent. This agent attempts to use said data on an application 2 without PAD and succeeds. It further attempts to use the data on an application 3 with PAD and fails

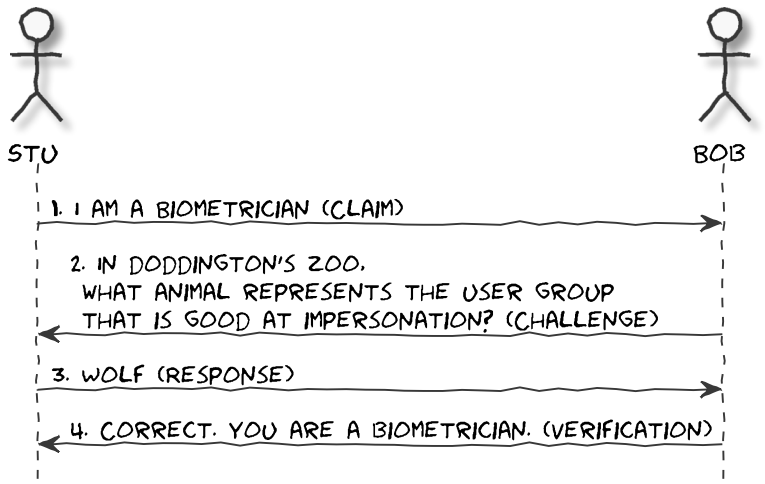

Fig. 2 Depicts a challenge-response protocol in the form of an interaction between a student Stu and a biometrician Bob. The claim is made in step 1, a challenge is issued in 2 and responded to in 3. Then, in 4, the authenticity of the initial claim has been verified, and Stu is thus a verified biometrician

generate responses without controlling said process. The space of possible questions in an oral exam would make it very unlikely that one would be able to guess the correct answer to each question, and learning every possible answer by heart would be very difficult without gaining some deeper underlying understanding (and thus passing the exam).

In certain digital signature algorithms, if no algorithmic and implementation flaws are found, guessing (or brute-forcing) a correct signature would take far longer than the time we have until the heat death of the universe. CAPTCHAs are another example where challenge-response protocols can be used to distinguish "robots" from humans, because there still exist easy-to-verify tasks that humans excel at relative to (machine learning) algorithms. Similarly, to beat challenge-response protocols in biometrics, a simple photograph of a face will no longer work if the biometric system checks for specific changes in light reflectance when a person is challenged to move their head.

3.2 Biometric challenge-response protocols

In biometric challenge-response protocols, the physiology or behavior of a user can be challenged to respond in a specific way in order to establish liveness of the presentation. For example, we expect pupil dilation as we shine a light in the user's eye; we expect changes to the skin's light reflectance when we ask a user to look in a different direction. We explicitly name this type biometric challenge-response protocols, as the processor of the challenge is a biometric *source*. Many other types of challenge-response protocols exist; some are also implemented in biometric systems as measures to protect the security and integrity of components of the system. These generally do not require involvement of the biometric source as a processor.

Figure 3 illustrates a model implementation of the essential components of a system that contains a biometric challenge-response protocol. This contains a challenge generator that is able to generate a biometric challenge (e.g., an instruction to move one's head) and dispatch it to both a response processor and a biometric source. A biometric source, the user of the system, is able to present its biometric data to a sensor and is able to process generated biometric challenges into biometric responses. A sensor is able to capture the biometric samples and responses, and finally, a response verifier is able to verify the validity of the response as a match to the challenge. The decision can be then propagated through components higher up in the hierarchy.

In a way, the link between the sensor and a specific biometric source constitutes a communication channel, across which information related to the authenticity of the source is transferred. A successful presentation attack is a compromise of the integrity of this channel. Biometric challenge-response protocols allow the introduction of a freshness component to the presentation of a biometric characteristic, thus promising mitigation of presentation attack threats at the source.

4 Formalism and categorization

We can now tie the concept of biometric challenge-response protocols into the formalism we introduced earlier. Further, we introduce the categorization for

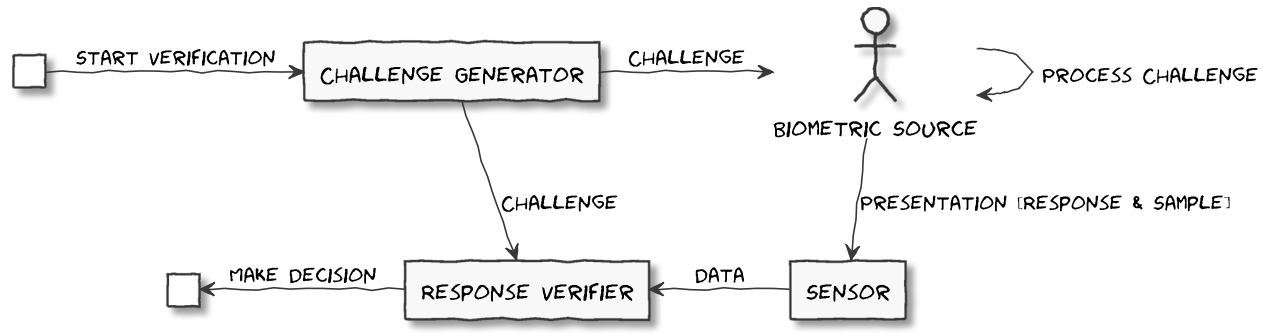

Fig. 3 A model biometric system with challenge response as it might be embedded in an existing biometric system. It includes a challenge generator (top-left), a biometric source (top-right), a sensor (bottom-right) and a response verifier (bottom-left)

biometric challenge-response protocols, which is independent of modality and whether its application lies in PAD or biometric recognition (BR). We show how this categorization contrasts against and improves upon the categorization as described in Table 3 of section 6.2.1 of ISO-30701-1.

4.1 Challenge-response protocol formalization extension

The formalism extension we suggest is relatively compact. We can include biometric challenge-response by saying that model M is now a generative model and is capable of generating challenges C; see Eq. 5.

$$M \rightarrow C \tag{5}$$

Source S then generates X during a biometric presentation in the context of the challenge C. This also has the property of reverting back to Eq. 1 if $C = \emptyset$.

$$S(C) \rightarrow X \tag{6}$$

The reason why the captured data X is not simply called a sample is now evident. It not only contains specific information related to the captured biometric instance but also information about the challenge, in the form of the response R, indicated by I_R. Equation 2 is extended to Eq. 7 to also include this response.

$$X = \{I_M, I_I, I_R, I_{\text{Other}}\} \tag{7}$$

The verification question of Eq. 4 is similarly extended to Eq. 8.

$$F(X, M, T, C) = D \tag{8}$$

The scoring function F thus now takes an additional challenge and considers X to contain the response to challenge C and using a model M transformed into a similarity score compared against threshold T, resulting in a decision D.

4.2 Challenge-response protocols and the weakest link

We can use the above extension to discuss challenge-response protocol's potential to provide additional protection to biometric applications, as was alluded to in the introduction. Like presented in an earlier section, the data X_1 that is still leaking from application 1 is retrieved. We assume an improved attack by the same malicious agent, where the agent has extended its previously created S_1^* with general modality information retrieved from either a PAD application leaking data, or more generally from research papers in the area. This creates S_1^{**}.

When the fake S_1^{**} is presented to the PAD-implementing application, the attack is likely to succeed, considering specific measures to include modality information I_M have now been taken, and the PAD application is no longer able to distinguish genuine from fake modalities. However, correctly implemented, we would expect an attack on the challenge-response protocol application to still be foiled.

S_1^* would be specific to C_1, whereas in the verification session, some challenge C_2 would be requested, resulting in a mismatch in the expected and received I_R.

We would want application 2, which has taken no special security precautions, to be secure against application 1's leaks as well. This is not something that happens by default, as we showed earlier. The crux will be to find a challenge-response protocol that makes identity verification decisions based on a combination of I_I and I_R, where ideally the I_I necessary is reduced and distorted by the user performing the response to challenge C_1.

4.3 Categorization

Table 3 of section 6.2.1 of ISO-30701-1 suggests a challenge-response categorization consisting of three categories: involuntary, voluntary and a "combination of something you know and are" (category 3). Involuntary

challenge-response protocols are considered those challenge-response protocols where responses are natural and not controllable by the subject. Voluntary, on the other hand, are based on "alive human cognition" and controlled action. Category 3 is also based on "alive human cognition," but requires specific individual biometric enrollment.

This categorization confounds the (natural) physiological vs behavioral responses, with "automated" vs "controlled" responses, assuming that physiological responses are always automated, and behavioral responses are always controlled. The former is not true, for example some people have voluntary control over their pupil dilation [9]. The latter is also not true, considering the existence, among many others, of the Stroop effect [10] or reflexive gaze changes. The voluntary vs. involuntary response categories seem more like a confounded dimension across which to categorize than categories of themselves.

Further, category 3 would in our formalism be a specific type of model M, that is in this category derived from specific individual biometric enrollments, as opposed to a "global" biometric enrollment. It is thus another dimension that is orthogonal to that of "involuntary vs. voluntary." We could imagine an involuntary individual challenge-response protocol, where a very detailed model is built of the way an individual's pupil dilates. We can also imagine a voluntary challenge-response protocol with a global/population-level model, e.g., through the aforementioned Stroop effect [10] or by employing general models of learning [11].

We thus designed our categorization with four concerns in mind. These are:

1. It should fit, but not overfit, current biometric challenge-response literature.
2. Dimensions of the categorization should be chosen such that they are mostly independent.
3. It should solve the confounds present in the ISO-30107 categorization.
4. It should be independent of modality-based categorization.

We believe the above criteria are reasonable, worthwhile, and cover what we want to achieve in a categorization.

We think the categorization described below meets these 4 criteria. In it, we introduce 5 categorization dimensions that allow for a clear and concise description of the biometric challenge-response protocol being introduced, while mostly abstracting away from specific biometric modalities or performance concerns. These dimensions are model (population to single subject), response type (physiological to behavioral), process (controlled to automated), noticeability (covert to overt), and skill required (unpracticed to practiced).

4.3.1 Model, population to single subject

Models (as defined in the formalism section) are used to verify whether the given response matches the challenge. The said models can be based on a characteristic a population possesses, e.g., pupil response to luminance level changes, or characteristics that smaller groups (down to an individual) possess, e.g., how a person's gait changes while stepping on an elevated surface.

The specificity of the model is closely related to whether the challenge-response protocol is usable for recognition or PAD. Although all biometric challenge-response protocols prove liveness of the presentation to a degree, a positive verification by a single subject model contains more identity information about the person than a population wide model would. In practice however, population wide models will be easier to build than single subject models, as single subject models need to be "trained" on the specific subject, thus making the enrollment phase more complex.

4.3.2 Response type, physiological to behavioral

Responses given as part of a presentation can broadly be classified as physiological responses, e.g., changes in skin conductance such as in [12], or behavioral, e.g., reflexive gaze changes in response to rapidly shifting stimuli [13].

This is one of two dimensions that are confounded in the ISO-30107 categorization and thus important to explicitly separate. The second dimension being the process dimension, as seen below.

4.3.3 Process, controlled to automated

Challenges can also vary in degree of conscious control someone has over starting and/or halting processing the challenge in the scope of the biometric system. Users can have full control over the response, e.g., speaking a challenge word [14], or the response can be mostly automated and/or reflexive, e.g., pupil dilation changes in response to certain stimuli [15].

Further, control can be given, even over automated/reflexive actions, by informing the user that a response is expected and can be denied in some way. In this case, even the "processing" of covert challenges such as ultrasonic sounds can be controlled by covering the lips or turning the head away. This has important security implications, as a user will in this way also be able to deny exhaustive recording of their response patterns and foil possible malicious actors in their attempts at attacking the system.

The other way around, the controlled process can become automated if a person practices them enough.

For example, the meaning of a word interfering with the task of speedily naming the color of that word, e.g., the red-colored word green. This effect is known as the Stroop effect [10] and is one of the most robust findings in experimental psychology. It works because it is near impossible to turn off the ability of reading and interpreting words.

4.3.4 Noticeability, covert to overt

Challenges can vary in degree in which they are noticeable by a person. This can range from fully covert, e.g., ultrasonic sounds causing fine vibrations in the lips while speaking [16], partially covert, e.g., determining ear canal structure by playing sounds into a person's ear [17], or very overt, e.g., in the patterns drawn with a finger on a smartphone [18].

Noticeability of a challenge can have a big impact on user experience (UX). On the one hand, covert methods allow for smooth User Interfaces (UIs) that do not ask too much of the user. On the other hand, this also provides malicious actors a method of stealthily gathering recordings of the person's response patterns, which is a real concern in combination with "automated" processing of challenges. Overt methods may therefore have a less smooth UI, but perhaps the (real) security benefits will serve to improve UX.

4.3.5 Skill required, unpracticed to practiced

The ability to respond to a given challenge can be "unpracticed," without practice prior to use of the biometric application. An example is to have subjects place on a vibrating plate that measures the change in vibration due to the hand-placement [19]. Other time, the security of the challenge-response protocol relies strongly on the user having had some form of practice during initial phases of application use. In [15], patterns in pupil dilation are used for verification, which arise from the presentation of a sequence of novel and non-novel pictures. It is possible to determine (non-)novelty, because users were presented with a sufficiently individualized selection of pictures during an enrollment phase. These would be considered "practiced" abilities.

4.4 Example

We can use the categorization to describe the challenge-response protocol proposed in [8]. In this protocol, a user is asked to look in specific directions in front of a camera. The algorithm is claimed to recognize both individual faces, as well as detect liveness by the expected way facial landmarks will move through 3D space. In terms of our categorization, it has a *population*-based model, as it simply looks at face landmarks and how they move through 3D space upon users' head movements. The instruction to move the head a certain direction requires a *behavioral* response. Head movements are a *controlled* process and instructions are very *overt*. For the general population, moving your head is a thing that can be done by anyone from a relatively young age and as such is an *unpracticed* skill. So in short, it is a challenge-response protocol that is population-based, behavioral, controlled, overt, and unpracticed.

5 Validation

In this section, we present quantitative arguments for why our categorization meets criteria 1 and 2 and qualitative arguments for why the categorization meets criteria 2, 3, and 4. The quantitative arguments are supported by a near-exhaustive survey of the contemporary biometric challenge-response protocol literature, which we will present as well.

5.1 Quantitative validation

5.1.1 Methods

The survey was conducted by searching for an initial seed of papers that mentioned in some capacity challenge-response and biometrics. Interactive biometrics is a term that is sometimes used as an alternative to challenge-response, and thus, these manuscripts were surveyed too. We collected and filtered through the manuscripts referenced by this initial seed and those that reference any of the seed manuscripts. In total, we collected 601 manuscripts, whittling them down to a total of 25 academic papers and 8 patents published between 2002 [20] and 2019; see Fig. 4 for an overview of the publication rate.

We classified each manuscript with the categorization and noted whether the manuscript was a paper or patent and the general sentiment the manuscript authors had regarding whether the challenge-response protocol was most applicable to PAD or biometric recognition (or a mix). See Table 2 for an overview of where each manuscript fell in this classification. See Fig. 5 for an overview of how manuscripts are spread across individual categories.

5.1.2 Criterion 1: Fit, but not overfit

For this criterion, we look at how surveys are spread across individual categories, as depicted in Fig. 5. There is a good spread across the "model" (population vs single-subject) category; however, other categories are somewhat slanted towards respectively behavioral (24 vs 6), controlled (23 vs 9), overt (28 vs 5), and unpracticed (26 vs 7).

5.1.3 Criterion 2: Independence of dimensions

Testing for independence of the categories gives Table 3. Each "first" category label was given the value 0, and

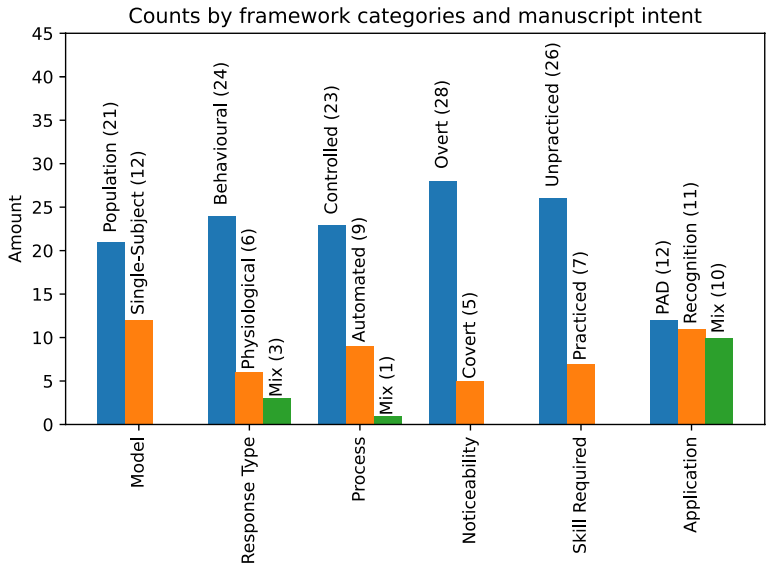

Fig. 4 Number of published biometric challenge-response protocol papers across the past 20 years

"second" category label was given the value 1, with mix being given the value 0.5. Pearson's r [44] was calculated for each category pair, as well as the p-value.

We find that the category "Skill required" (practiced to unpracticed) is not correlated to other categories. There is however a high correlation between "Model" and "Application" ($r = 0.56; p < 0.01$) and "Response type" and "Process" ($r = 0.64; p < 0.01$). The above results suggest that although there is some dependence between categories; overall, the independence of the dimensions is very good.

5.2 Qualitative validation

5.2.1 Criterion 2: Independence of dimensions

A more qualitative argument for the independence of categorization dimensions is to select a number of articles that are very far apart in the space spun by these dimensions. This in turn should result in articles that describe protocols that are very different in kind. What follows is a discussion of 4 articles [6, 12, 15, 23]. They show the diversity and creativity in types of challenge-response protocols that are currently being researched (and used), and how this diversity is represented in the space spun by our categorization.

5.2.2 Pupil light reflex [6]: single-subject model, physiological, automated process, overt, unpracticed

This article proposes a multi-modal system that fuses a score based on static features of the iris with a score based on dynamic features of an individual's pupil light reflex (PLR) elicited by a short light flash. The dynamic features come from a challenge-response protocol where the light flash is the challenge, and the PLR is the response. The model used is a single-subject model. The response relies on a physiological reaction to the light flash, which is also very overt. The reflex itself is automated and unpracticed.

Possible future extensions on this challenge-response protocol are expanding the space of challenges by varying the duration and the intensity of the flash, as well as the number of flashes. A very dim, singular short flash will not elicit the same response as a very bright multitude of longer flashes. Further, it could be extended towards a practiced, controlled challenge-response protocol, meaning a user would be allowed to close their eyes for the duration of some of the flashes, allowing the encoding of a password-like secret into the protocol.

5.2.3 Pulse-response biometrics [12]: single-subject model, physiological, controlled process, covert, unpracticed

This paper suggests a physical enhancement to a PIN-based authentication scheme. The authors suggest constructing ATM's in which the PIN-pad as well as a small flat surface are made of some conducting material. The user is asked to put one hand on the flat surface and to enter their PIN with their other hand. When the finger touches a button on the PIN-pad, a low voltage pulse signal (the challenge) can be applied to the finger and then measured in the palm of the other hand, which will be modulated by the user's specific physiology as it goes through their finger, hands, arms, and torso (the response).

Table 2 Table of manuscripts describing biometric challenge-response protocols. The columns represent where the manuscripts fall on the 5 categories, as well as the type of manuscript and PAD or recognition intent of the described protocol

Authors	Ref	(P)opulation to (S)ingle Subject	(B)ehavioral to (P)hysiological	(A)utomated to (C) ontrolled	(O)vert to (C) overt	(U)npracticed to (P)racticed	Manuscript Type	(R) ecognition to (P)AD
Simons et al. (2014)	[21]	P	B	A	C	U	Paper	P
Jakobsson (2015)	[22]	P	B	C	O	P	Paper	Mix
Frischholz and Werner (2003)	[8]	P	B	C	O	U	Paper	P
Ganesh et al. (2017)	[23]	P	B	C	O	U	Paper	P
De Marsico et al. (2012)	[24]	P	B	C	O	U	Paper	P
Saad and Moustafa (2014)	[25]	P	B	C	O	U	Paper	P
Shen et al. (2018)	[26]	P	B	C	O	U	Paper	P
Ali et al. (2012)	[27]	P	B	C	O	U	Paper	P
Li et al. (2016)	[28]	P	B	Mix	O	U	Paper	R
Zhang et al. (2018)	[29]	S	B	A	C	U	Paper	R
Gong et al. (2016)	[30]	S	B	C	C	P	Paper	Mix
Tan et al. (2018)	[16]	S	P	A	C	U	Paper	Mix
Martinovic et al. (2014)	[12]	S	P	C	C	U	Paper	R
Sluganovic (2016)	[13]	S	B	A	O	U	Paper	R
Bhardwaj (2017)	[31]	S	B	C	O	P	Paper	P
Burgbacher et al. (2014)	[18]	S	B	C	O	P	Paper	R
Ma et al. (2018)	[32]	S	B	C	O	P	Paper	R
Skerpac (2002)	[20]	S	B	C	O	U	Paper	Mix
Tian et al. (2017)	[33]	S	B	C	O	U	Paper	Mix
Sae-Bae et al. (2012)	[34]	S	B	C	O	U	Paper	R
Chauhan et al. (2017)	[35]	S	Mix	C	O	U	Paper	R
Sae-Bae and Jakobsson (2014)	[36]	S	Mix	C	O	U	Paper	R
Blanchard et al. (2019)	[15]	S	P	A	O	P	Paper	R
Li et al. (2019)	[19]	S	P	A	O	U	Paper	Mix
Yano et al. (2012)	[6]	S	P	A	O	U	Paper	P
Gao et al. (2019)	[17]	S	P	A	O	U	Paper	R
Bhaskaran (2011)	[37]	P	B	C	O	U	Patent	Mix
Chaudhury (2014)	[38]	P	B	C	O	U	Patent	P
Lindemann (2018)	[39]	P	Mix	A	O	U	Patent	P
Roos (2014)	[40]	S	B	C	O	P	Patent	P
Pearson and Contolini (2005)	[41]	S	B	C	O	U	Patent	Mix
Roblek and Sharifi (2014)	[42]	S	B	C	O	U	Patent	Mix
Skerpac (2014)	[43]	S	B	C	O	U	Patent	Mix

A single-subject model can be used to distinguish the physiological pulse-responses of individuals. The signal is itself covert, as the user of the system should not be feel said pulse. It is a controlled process in the sense that users can choose not to have the pulse go through their body, and it does not take any previous practice to make use of this challenge-response protocol.

This method is interesting as it combines a physiological response with a covert challenge. Although the manuscript does not describe a wide range of challenges/pulses being applied, it would be easy to imagine a set-up where each applied pulse is different from all earlier pulses.

5.2.4 Gesture-based rotary dial for phone unlocking [23]: population model, behavioral, controlled process, overt, unpracticed

The authors describe a smartphone-based PIN-entry pad styled like an old rotary phone where the PIN can be entered by sliding the finger from digit to digit. This generates both the PIN as a secret, as well as behavioral data, specifically touch pressure, that is used to

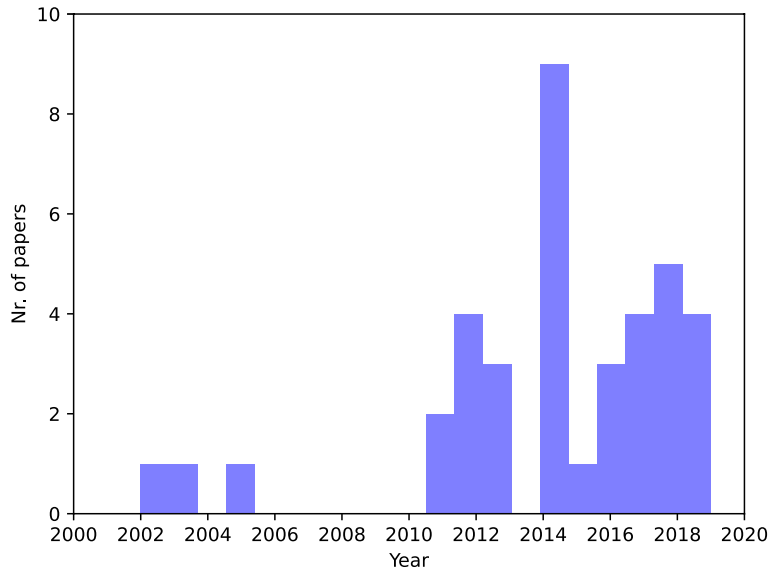

Fig. 5 Spread of papers across the different categories.

Table 3 Table of correlations, Pearson's *r*, between individual categories on surveyed manuscripts. Stars represent significant *p*-values. One star: $p < 0.05$, two stars: $p < 0.01$

	Model	Response type	Process	Noticeability	Skill required	Application
Model	**1.00	*0.36	0.14	0.14	0.24	**0.58
Response type		**1.00	**0.58	0.19	-0.11	0.25
Process			**1.00	0.30	-0.17	0.15
Noticeability				**1.00	-0.01	0.12
Skill required					**1.00	0.11
Application						**1.00

authenticate the user's identity. A user's PIN is consistent, but the ordering of the PIN pad numbers (the challenge) is changed for each PIN-pad session. There are thus 10! unique PIN pads to be created. The feature that are used are claimed to be independent of the ordering of the pad, and thus, the response only consists of entering the correct PIN in this session-unique pin pad. We therefore consider the model to be of population level, as a successful PIN entry (whatever the PIN may be) constitutes a successful behavioral and controlled response. No real prior practice is needed before using this system.

5.2.5 Image familiarity as your password [15]: single-subject model, physiological, automated process, overt, practiced

In this method, users are shown a series of pictures in the enrollment phase [15]. During verification, users are shown a selection of familiar pictures from the enrollment phase, as well as unfamiliar pictures in rapid succession (the challenge). Pupil dilation is used to measure "novelty" [45] of these images. Image novelty in sequence forms a bit-pattern (the response) that can then be matched against the novelty bit-pattern the system expects for the user.

Every user is shown a different set of images, from which a single-subject model is derived. The response is physiological, as it relates to pupil dilation. The images themselves are very overt, but the processing is done automatically. Practice is required as familiarity with a selection of pictures is needed to successfully complete the verification phase.

5.2.6 Criterion 3: Solve confounds in ISO-30107 categorization

In our eyes, the ISO-30107 categorization of voluntary, involuntary, and a type of single-subject model confounds the dimensions of physiological vs behavioral responses with a controlled vs automated type of

processing. Further, category 3 seems to suggest a dimension on its own, which can be applied to both voluntary and involuntary processes. These confounds are solved by introducing all three as their own separate categorization dimensions, which we have done in our categorization.

5.2.7 Criterion 4: Independence of modality-based distinctions

The final criterion is that a given categorization should give no indications of what modality is being employed in, nor if the author's intention was originally a PAD application or BR. The former is reasonably achieved we believe, although limitations exist.

For example, if the modality is DNA, and if we can find some way to challenge this modality, the response type will certainly not be behavioral. On the other hand, a categorization such as s*ingle-subject model, physiological*, *automated process*, *overt*, and *practiced* would not, to us, suggest a specific modality but would suggest important features of the application's UI/UX regarding a required practice and enrollment phase, as well as the security implications of using "automated" response processes.

6 Discussion

The formalism and categorization we introduced provide an interesting window into the issues raised in the introduction of this manuscript, namely the "weakest link problem" in the context of biometric database leaks. Specifically, we discuss how one of our categorization's dimension in particular maps onto the "right kind" of challenge-response protocols to provide system "insulation." Further, we discuss some of the validation results we presented for the categorization.

6.1 Mitigating the w*eakest link problem*

Big biometric data leaks, such as the leak of personal identifying information combined with high resolution scans of fingerprints and irises from the Biostar 2 database [1] and the leak of 1+ million DNA profiles from the GEDmatch database [46], are becoming a regular occurrence. It is unfortunate that good security practices related to databases and permission handling cannot be universally expected and that applications of the data beyond biometric authentication cause these data to be sometimes stored in a less-than-ideal format (from an authentication point of view), e.g., forensic applications sometimes require partial fingerprint matching, which forces fingerprints to be stored in a "raw" photo-quality format.

These leaks provide malicious actors with the data to start presentation attacks such as pictured in Fig. 1. The "shared" resource pool, in an individual biometric source, from which these data are drawn, makes such attacks potentially very effective, such that all biometric applications share a *weakest link problem*.

We claim that the right kind of biometric challenge-response protocols provide a stronger defensive strategy than "simple" PAD methods can muster. The type of PAD methods that are not challenge-response protocols are necessarily limited to passive collection of data. Challenge-response protocol methods in general are able to elicit information through a challenge (that can be made verification-session-dependent) from a process in the biometric source that should be difficult to replicate. In contrast, non-CRP PAD methods likely have no such session-dependent information. Successful presentation attacks should thus already be harder through the use of challenge-response mechanisms.

This advantage becomes further enhanced if one considers response verification models that are s*ingle-subject*. Data X, as introduced in the formalism, captured by the biometric sensor, in a sufficiently strong challenge-response protocol, comprises both information from the biometric characteristic of the source S *and* the challenge C. It should only be possible to verify X if the original C, as well as the model M all match with the captured X. Replaying X at a later point in time to a system with M will make C, and the resulting I_R no longer match the I_R present in X, similarly for replaying X on a different system's M'. Replaying X, directly on the sensor or indirectly through an artifact (fingerprint mold, photo of face), that *does not* use a biometric challenge-response protocol, will ideally make the verification on that system fail, due to the "noise" added by the response to C. Alternately, replaying an X' captured from a M' that does not implement a biometric challenge-response protocol will fail because (1) no information relevant to C will be present and (2) X' is only a representation of S, and not S itself.

Further, in a single-subject model M, the source S is considered to be individualized, and the derived M is a model of the specific individual, not of a population. A verification of X based on M and C can therefore be claimed to answer the question: *Is it really this person?*. The interactions described above also indicate that were a leak to happen in a system implementing biometric challenge-response protocol, other systems (with and without biometric challenge-response protocol) do not suffer the consequences of security lapses as much. More importantly, systems implementing the said biometric challenge-response protocol are insulated from leaks happening in other systems as well. So, although the weakest link problem will still exist as long as a chain of conventional biometric systems exist, s*ingle-subject* biometric challenge-response protocols are a step towards unlinking the chain without breaking it.

6.2 Categorization validation

To reiterate, our categorization consists of the following category categories: model, response type, process, noticeability, and skill required. We tested the fit of these 5 categories, as well as the "sentiment" of the authors in regards to whether their method is a PAD or recognition method, to the current state of the literature. We found that there is a slight-to-medium bias to one category in each dimension except for model, which is fairly evenly distributed. Further, we tested the interdependence of each category and found high correlations between response type and noticeability, as well as response type and process. Further model and application correlated highly, indicating that the model dimension mostly maps on to what authors believe their method can be used for, which in a way validates the move of abstracting way from PAD and recognition, while still retaining the same discriminatory power.

In any case, the two categories that correlate highly with response type are not necessarily an issue with the categorization but can be explained by biases in the current state of the literature. From the analysis, we find that physiological response types are often combined with automated processes, whereas behavioral response types are combined with controlled processes and overt challenges.

These categories do not necessarily have to always go together. We can imagine physiological responses such as the pupil reflex that can be controlled by closing the eye-lids. If the challenge-response protocol accommodates for this behavior, we would categorize the said process to be controlled. There are also plenty of opportunities to do covert challenges that elicit behavioral responses, such as by building on work by [30] and [39].

An interesting experiment is to take existing challenge-response protocols and change their categorization in one of the five dimensions to use as a starting point for new challenge-response protocols, e.g., we can start with the image familiarity [15] method and imagine it using a population-wide model, rather than a single-subject model. Enrollment would then consist of a set of images that is the same for each user, which they are trained on. During verification, images from this set and novel images would be mixed; the correct bit pattern of non-novel and novel images would say something about the liveness of the user. One could imagine a face recognition system being augmented with this "population-wide" liveness detection method.

7 Conclusion

Presentation attack detection (PAD) concerns the detection of attacks on a biometric recognition system. Biometric recognition suffers from a *weakest link problem*, where a leak of one biometric application's database forms the basis from which a malicious agent can mount attacks on *other* biometric authentication applications. This is a serious security concern for both individuals using these applications and the application administrators. In this paper, we suggest that the "right kind" of challenge-response protocols can provide additional mitigation of the weakest link problem.

We introduced a set of concepts, a formalism, and categorization for biometric challenge-response protocols. The formalism expresses the abstract similarities between PAD and recognition and then extends it to include biometric challenge-response protocols. The categorization extends the categorization as proposed in ISO-30107-1 and allows us to structurally understand biometric challenge-response protocols through a model comprising 5 categories.

We showed the categorization fits the current state of the literature reasonably well, and that each category is reasonably independent. Although its non-perfect fit and independence can be considered a weak point as well. We used the biometric challenge-response protocol formalization to show how s*ingle-subject* category challenge-response protocols are the "right" kind of challenge-response protocols. They have the desirable feature that leaks from these individual applications should have a significantly lowered impact on other applications, with and without PAD.

We believe biometric challenge-response protocols are important for the security of biometric applications. Contemporary research into biometric challenge-response protocols is unfortunately very fragmented. Our intent and hope for the formalism and categorization is therefore to focus and give direction to further research into biometric challenge-response protocols, so that the impact of future data leaks may become less and less.

Abbreviations

TOR: The Onion Router; BCRP: Biometric challenge-response protocol; CR: Challenge-response; CRP: Challenge-response protocol; ISO: International Organization for Standardization; PAD: Presentation attack detection; PLR: Pupil light reflex; UI: User interface; UX: User experience.

Acknowledgements

We thank Bram Kolkman for assistance in designing several figures. We thank Léon Melis for proofreading drafts of the paper.

Authors' contributions

Author E. Haasnoot wrote the paper. Co-authors L.J. Spreeuwers and R.N.J. Veldhuis are E. Haasnoot's PhD advisors and worked on the paper in that capacity. All authors read and approved the final manuscript.

Funding
Not applicable

Availability of data and materials
Not applicable

Declarations

Ethics approval and consent to participate
Not applicable

Consent for publication
Not applicable

Competing interests
The authors declare that they have no competing interests.

Received: 25 September 2020 Accepted: 16 August 2022
Published online: 05 September 2022

References

1. G. Fawkes, vpnMentor, Report: Data breach in biometric security platform affecting millions of users. (2019). https://www.vpnmentor.com/blog/report-biostar2-leak/. Accessed 14 Aug 2019
2. R. Singh, R. Khaira, Rs 500, 10 minutes, and you have access to [a] billion aadhaar details. (2019). https://www.tribuneindia.com/news/archive/nation/rs-500-10-minutes-and-you-have-access-to-billion-aadhaar-details-523361. Accessed 05 Jan 2018
3. R. Dingledine, N. Mathewson, P. Syverson, Tor: The second-generation onion router. Technical report (NavalResearch Lab, Washington DC, 2004)
4. S. Nakamoto, Bitcoin: a peer-to-peer electronic cash system. Whitepaper (2008). https://bitcoin.org/bitcoin.pdf
5. L. Matney, Sam Altman's Worldcoin wants to scan eyeballs in exchange for crypto (2021). https://techcrunch.com/2021/10/21/sam-altmans-worldcoin-wants-to-scan-every-humans-eyeball-and-give-them-crypto-in-exchange/. Accessed 21 Oct 2021
6. V. Yano, A. Zimmer, L.L. Ling, in Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012). Multimodal biometric authentication based on iris pattern and pupil light reflex (IEEE, New York, 2012), pp. 2857–2860
7. International Organization for Standardization, vol. 2016 (International Organization for Standardization, Geneva, 2016)
8. R.W. Frischholz, A. Werner, in 2003 IEEE International SOI Conference. Proceedings (Cat. No. 03CH37443). Avoiding replay-attacks in a face recognition system using head-pose estimation (IEEE, New York, 2003), pp. 234–235
9. L.V. Eberhardt, G. Grön, M. Ulrich, A. Huckauf, C. Strauch, Direct voluntary control of pupil constriction and dilation: exploratory evidence from pupillometry, optometry, skin conductance, perception, and functional MRI. Int. J. Psychophysiol. **168**, 33–42 (2021)
10. J.R. Stroop, Studies of interference in serial verbal reactions. J. Exp. Psychol. **18**(6), 643 (1935)
11. A. Khodabakhsh, E. Haasnoot, P. Bours, in 2018 International Conference of the Biometrics Special Interest Group (BIOSIG). Predicted templates: learning-curve based template projection for keystroke dynamics (IEEE, New York, 2018), pp. 1–5
12. I. Martinovic, K. Rasmussen, M. Roeschlin, G. Tsudik, Authentication using pulse-response biometrics. Commun. ACM. **60**, 108–115 (2017)
13. I. Sluganovic, M. Roeschlin, K.B. Rasmussen, I. Martinovic, in Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. Using reflexive eye movements for fast challenge-response authentication (ACM, New York, 2016), pp. 1056–1067
14. R. Johnson, T.E. Boult, W.J. Scheirer, in 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS). Voice authentication using short phrases: examining accuracy, security and privacy issues (IEEE, New York, 2013), pp. 1–8
15. N.K. Blanchard, S. Kachanovich, T. Selker, F. Waligorski, Reflexive memory authenticator: a proposal for effortless renewable biometrics. Emerging technologies for authorization and authentication, **11967**, 104–121 (2020)
16. J. Tan, X. Wang, C..-T.. Nguyen, Y. Shi, SilentKey: A new authentication framework through ultrasonic-based lip reading. Proc. ACM Interact. Mob. Wearable Ubiquit. Technol. **2**(1), 36 (2018)
17. Y. Gao, W. Wang, V.V. Phoha, W. Sun, Z. Jin, EarEcho: Using Ear Canal Echo for Wearable Authentication. Proc. ACM Interact. Mob. Wearable Ubiquit. Technol. **3**(3), 1–24 (2019)
18. U. Burgbacher, M. Prätorius, K. Hinrichs, in 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC). A behavioral biometric challenge and response approach to user authentication on smartphones (IEEE, New York, 2014), pp. 3328–3335
19. J. Li, K. Fawaz, Y. Kim, in Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security. Velody: nonlinear vibration challenge-response for resilient user authentication (ACM, New York, 2019). pp. 1201–1213
20. V. Skerpac, N-dimensional biometric security system (2002). US Patent App. 10/062,799
21. S. Simons, J. Zhou, Y. Liao, L. Bradway, M. Aguilar, P.M. Connolly, Cognitive biometrics using mouse perturbation (2014). US Patent App. 14/011,351
22. B.M. Jakobsson, Systems and methods for authenticating a user based on a biometric model associated with the user (2015). US Patent 9,203,835
23. S.M. Ganesh, P. Vijayakumar, L.J. Deborah, in 2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM). A secure gesture based authentication scheme to unlock the smartphones (IEEE, New York, 2017), pp. 153–158
24. M. De Marsico, M. Nappi, D. Riccio J.-L. Dugelay, in 2012 5th IAPR International Conference on Biometrics (ICB). Moving face spoofing detection via 3D projective invariants (IEEE, New York, 2012), pp. 73–78
25. A. Saad, M. Moustafa, in Proceedings of the International Conference on Artificial Intelligence and Pattern Recognition. Interactive versus Passive 2D Face Spoofing Detection (ACM, New York, 2014)
26. M. Shen, Y. Wei, Z. Liao, L. Zhu, Iritrack: Face presentation attack detection using iris tracking. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. **5**(2), 78–17821 (2021). https://doi.org/10.1145/3463515
27. A. Ali, F. Deravi, S. Hoque, in 2012 Third International Conference on Emerging Security Technologies. Liveness detection using gaze collinearity (IEEE, New York, 2012), pp. 62–65
28. S. Li, A. Ashok, Y. Zhang, C. Xu, J. Lindqvist, M. Gruteser, in 2016 IEEE International Conference on Pervasive Computing and Communications (PerCom). Whose move is it anyway? Authenticating smart wearable devices using unique head movement patterns (IEEE, New York, 2016), pp. 1–9
29. Y. Zhang, W. Hu, W. Xu, C.T. Chou, J. Hu, Continuous authentication using eye movement response of implicit visual stimuli. Proc. ACM Interact. Mob. Wearable Ubiquit. Technol. **1**(4), 177 (2018)
30. N.Z. Gong, M. Payer, R. Moazzezi, M. Frank, in Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security. Forgery-resistant touch-based authentication on mobile devices (ACM, New York, 2016), pp. 499–510
31. I. Bhardwaj, N.D. Londhe, S.K. Kopparapu, Study of imposter attacks on novel fingerprint dynamics based verification system. IEEE Access. **5**, 595–606 (2016)
32. Z. Ma, X. Wang, R. Ma, Z. Wang, J. Ma, Integrating gaze tracking and head-motion prediction for mobile device authentication: a proof of concept. Sensors. **18**(9), 2894 (2018)
33. J. Tian, Y. Cao, W. Xu, S. Wang, Challenge-response authentication using in-air handwriting style verification. IEEE Transactions on Dependable and Secure Computing **17**(1), 51–64 (2020)
34. N. Sae-Bae, K. Ahmed, K. Isbister, N. Memon, in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Biometric-rich gestures: a novel approach to authentication on multi-touch devices (ACM, New York, 2012), pp. 977–986
35. J. Chauhan, Y. Hu, S. Seneviratne, A. Misra, A. Seneviratne, Y. Lee, in Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services. BreathPrint: Breathing acoustics-based user authentication (ACM, New York, 2017), pp. 278–291
36. N. Sae-Bae, M. Jakobsson, in Proceedings of the 15th Workshop on Mobile Computing Systems and Applications. Hand authentication on multi-touch tablets (ACM, New York, 2014), p. 8

37. S. Bhaskaran, Integrated voice biometrics cloud security gateway. (2016). US Patent 9,412,381
38. K. Chaudhury, A. Devarasetty, Liveness detection (2014). US Patent 8,856,541
39. R. Lindemann, System and method for eye tracking during authentication (2018). US Patent 9,898,596
40. D. Roos, System and methods for personal identification number authentication and verification (2018). US Patent 10,049,197
41. S. Pearson, M. Contolini, System and method for portable authentication (2005). US Patent App. 10/859,487
42. D. Roblek, M. Sharifi, Segment-based speaker verification using dynamically generated phrases (2014). US Patent 8,812,320
43. V. Skerpac, Dynamic pass phrase security system (dpss) (2014). US Patent 8,812,319
44. K. Pearson, VII. Note on regression and inheritance in the case of two parents. Proc. R. Soc. Lond. **58**(347–352), 240–242 (1895)
45. M. Naber, S. Frässle, U. Rutishauser, W. Einhäuser, Pupil size signals novelty and predicts later retrieval success for declarative memories of natural scenes. J Vis. **13**(2), 11 (2013)
46. H. Murphy, Why a data breach at a genealogy site has privacy experts worried (2020). https://www.nytimes.com/2020/08/01/technology/gedmatch-breach-privacy.html. Accessed 01 Aug 2020

Publisher's Note