Principles of Authentication

Sean Peisert
UC Davis and Berkeley Lab
California, USA
peisert@cs.ucdavis.edu

Ed Talbot
UC Davis
California, USA
edward.talbot@gmail.com

Tom Kroeger
Sandia National Laboratories
Calfornia, USA
tmkroeg@sandia.gov

ABSTRACT

In the real world we do authentication hundreds of times a day with little effort and strong confidence. We believe that the digital world can and should catch up. The focus of this paper is about authentication for critical applications. Specifically, it is about the fundamentals for evaluating whether or not someone is who they say they are by using combinations of multiple meaningful and measurable input factors. We present a "gold standard" for authentication that builds from what we naturally and effortlessly do everyday in a face-to-face meeting. We also consider how such authentication systems can enable resilience to users under duress. This work differs from much of the other work in authentication first by focusing on authentication techniques that provide meaningful measures of confidence in identity and also by using a multifaceted approach that comprehensively integrates multiple factors into a continuous authentication system, without adding burdensome overhead to users.

Categories and Subject Descriptors

D.4.6 [**Operating Systems**]: Security and Protection; K.6.5 [**Management of Computing and Information Systems**]: Security and Protection; H.1 [**Information Systems**]: Models and Principles

General Terms

Design, Management, Security

Keywords

Authentication; biometrics; continuous authentication; duress codes; multi-dimensional inputs; principles of authentication

NSPW'13, September 9–12, 2013, Banff, AB, Canada.

ACM 978-1-4503-2582-0/13/09 ...$15.00.
http://dx.doi.org/10.1145/2535813.2535819.

"Who are you?" said the Caterpillar.

Alice replied, rather shyly, "I—I hardly know, sir, just at present—at least I know who I was when I got up this morning, but I think I must have changed several times since then."

—Lewis Carroll, *Alice's Adventures in Wonderland* (1865)

1. INTRODUCTION

Systems can be measurably secured against attacks on availability, confidentiality, and integrity using clean-slate, ground-up techniques involving combinations of formal verification and both technological and "human" Byzantine fault tolerance [40]. However, all such systems—even high-assurance, critical systems (e.g., nuclear command and control [3, §13.2.2]) that use diverse, independent redundancy from the atoms composing a system's transistors to the humans operating the system—require measurably strong validation and verification of human identity. In this paper,we discuss solutions for authenticating in critical environments. We also believe that many elements of our solutions could be applicable elsewhere in non-critical environments, and we seek to provide for mechanisms that can allow this.

Authentication, sometimes called *origin integrity*, is a means of measuring the degree of trust that one can have that the source of data is who it purports to be [6, §1.1.2]. Two concepts are implicit in origin integrity: 1) the source of the data is indeed who it claims to be (identity integrity), and 2) the data is a faithful reproduction from that source (data integrity). This paper is concerned mostly with the former. Humans have authenticated to each other throughout history. Some of the time, that authentication has been between two people physically near each other. Sometimes two people cannot be near each other, however, or may not know each other's appearance, so alternate means have been used, such as the impressions of signet rings in wax, secret handshakes, or passwords. The reality of these latter techniques is that they often failed to correctly validate identity. Today, we still authenticate to each other by recognizing one another when we are in close physical proximity.

But, as it has always been, there are times when authentication over a distance is required. And indeed, just as in earlier times, the techniques that we use today often fail as well. We assert that this is because most of the techniques currently used for authentication over a distance do not *measurably* provide a meaningful degree of trust. Knowing a password may indicate nothing more than that the password has not been guessed, and possessing an one-time

password token may indicate nothing more than it has been stolen. And it is impossible to measure the risk of either.

Thus, the focus of this paper is twofold: first, it is about methods for using data input factors across multiple dimensions that provide meaningful and measurable confidence about whether or not someone is who they say they are, and that they are not, for example, performing a masquerade attack [31] by presenting stolen, forged, duplicated, guessed, or mimicked credentials. This paper is also about making sure that the authentication is intentional and not, for example, coerced. These two concepts—authentication and intent—form the basis for definitive command and control of critical systems [3, §13.2.2]. We briefly touch on the problem of securing systems against people who are already trusted and who then decide to do something malicious (e.g., "insiders" [8]) but primarily leave system design, formal methods, and fault tolerance as defenses to address such threats [40]. We conclude the paper by presenting four "principles of authentication" that we assert that systems must adhere to in order to capture the meaningful and measurable elements that make the current *de facto* standard of in-person, human-to-human function well.

2. BACKGROUND

In the physical world, we perform authentication trivially. For example, humans may see someone whose face they recognize in a context they know, such as their workplace. If they see someone who they do not know in a sensitive area, then suspicion is raised. If they see someone they know in a place they do not expect to see them (e.g., a bank in the Cayman Islands), their suspicion may also be raised. It is this trivial, intuitive function that humans perform naturally. But it is difficult to be in a situation in which humans can always access a particular set of computational resources by authenticating themselves to another human. So we rely on computers to perform this authentication function. We note that the Turing Test [48] represents a lighter class of the authentication problem and has demonstrated the challenges in simply identifying that someone is a human. Identifying that someone is a *specific* person is a much harder problem that humans are far better at performing.

Authentication to a computer today typically relies on one or more (called *multifactor authentication*) of the following: something you know, such as passwords; something you have, such as one-time password hardware tokens; and/or something you are, such as some form of biometric [6, §12]. Individual authentication techniques typically include traditional passwords, graphical and video passwords, cognitive (e.g., word association) passwords, "tokens" (paper, hardware, etc..), and biometrics. Individually, these authentication techniques have a wide range of user effort, scalability, learning curve, accessibility, and resilience to theft, observation, and guessing [9, 10, 17, 19, 42]. Multi-factor authentication [20] and multi-step authentication are solid steps toward improving many of the key qualities for robust authentication. [49]. However, these criteria are likely to all be relevant to traditional computing, but traditional computing has a vastly different set of criteria than our question of validating identity in critical environments using inputs that provide *meaningful* and *measurable* degrees of confidence.

The state of digital authentication is not very impressive. Even for non-critical environments, one need only look at the number of times that unauthorized users access resources they should not have access to, or authorized users are accidentally granted access to resources that they should not, to know that authentication techniques currently in use are grossly inadequate. Moreover, as we discussed earlier, they do not actually address the problem: helping to verify the identity and intent of a individual. For example, long, complicated passwords are useful defenses against the threat of someone stealing a file containing password hashes. However, they are no longer relevant when the system being authenticated to locks out or throttles passwords after sufficient incorrect guesses [5]. Moreover, Person A knowing a password does not give Person B any valid means of determining whether Person A is who they claim to be. Even combinations of different classes of techniques—e.g., combining a password with an one-time password token and word association still do not address this problem [10].

And yet, since humans perform authentication effortlessly, intuitively, and naturally we tend to take it for granted. We instinctively believe authentication is strong even when in reality, it is supported by the flimsiest evidence. The continued domination of passwords over other methods of authentication is an example of such instinctive belief.

3. GOALS AND USE CASES

We assert that physical, in-person interaction between two people who recognize each other is the "*de facto* standard" for authentication. It is *not* perfect. Human memories can fail and people's appearance can become less recognizable (e.g., due to age) or masqueraded (e.g., via surgery). However, human-to-human is based on more than a snapshot of appearance. There are gestures, gaits, and other patterns that provide substantially more input. Studies have shown that words only convey about 7% of the message content in a face-to-face exchange [32]. The remaining 93% is conveyed through tone of voice and body language.

The strength of authentication via passwords is very limited [32]. Password authentication to computers was developed in an environment where every bit, flop, and memory cell was precious. Technology today enables far richer authentication protocols. Therefore, the standard for comparison of authentication schemes should be the canonical face-to-face encounter between humans and not the existing password paradigm. Putting this situation in context, we propose a simple model for authentication that enables consideration of authentication in technology-enabled situations. This model includes face-to-face communication, communication through a pipe (or conduit), and communication with delay (or storage). These modes of human interaction are developed in the sections below:

Face-to-Face communication is the base case of interaction between two or more people. But such communication is limited because it requires both geographic and temporal synchronization. Technology enables asynchronous communication or communication over a distance but that convenience comes at a cost, i.e., eroded confidence of identity. Passwords are the de facto means of mitigating such erosion. However, as we describe below, the judicious application of technology in the other modes of communication can enable dramatic improvements in confidence of identity through increased natural cueing. The objective in all other modes of human interaction requiring authentication should strive to provide measures of identity that are as strong as in face-to-face interaction.

A *pipe (or conduit)* enables human interaction without requiring geographic synchronization. Videoconferencing and telephone conversations are examples of such "pipe-enabled" interaction. Both video and telephone enable authentication via tone of voice. Videoconferencing provides richer authentication because it enables the users to view geographic context and, perhaps even more importantly, body language.

Conversely, a "While You Were Out" sticky note is an example of human interaction that requires geographic synchronization without requiring temporal synchronization.

Delay (or storage) enables human interaction in the absence of both geographic and temporal synchronization. E-mail is an example of such delay-enabled interaction (as are social networking sites such as Facebook). The fact that e-mail uses a server to store messages requires that users authenticate themselves to a machine to use the system. The machine becomes an intermediary between humans and the interaction is now mediated by authentication that the machine is capable of handling (i.e., passwords). This richer authentication of e-mail can be enabled by eliminating the server in favor of a peer-to-peer e-mail paradigm because a sysadmin for a central mail service cannot possibly take the time to assess the rich authentication inputs for every single email routed through the system, even in relatively low-volume environments. In contrast, something more akin to peer-to-peer or a hybrid between peer-to-peer and centralized mail could make this tractable.

Applying our model for human interaction to scenarios in which geographic or temporal synchronization are not possible suggests the use of *richer* authentication techniques—techniques that communicate substantially more data about a person *relevant* to the authentication process—than those that are commonly used today. Rich authentication demands that users reveal significant information about themselves. This is a natural outcome in face-to-face encounters, but it is suspect in technology-enabled interaction because current authentication technology does not typically demand such revelation. As such, while our focus is on critical systems and we therefore largely ignore privacy, one could imagine applying these techniques in non-critical environments. In such cases, privacy considerations may have more weight. For simple transactions such as browsing the news on a website, users may choose to reveal very little about themselves and the news service may require only modest user information. As the transaction becomes more important (e.g., banking or national security) the user may be required to reveal more information about themselves. Trust negotiation approaches have been developed to facilitate such interaction [28].

Certainly the highest level of authentication also demands that humans be the ultimate arbiters of authentication. Machines can provide information to facilitate such arbitration but *the authentication decision must rest with solely with humans*. It is using this notion that we discuss our goals and assumptions:

Goals.

Authentication should be measurably precise. It should never be unintentional or accidental [45]. It must not be sharable, or vulnerable to loss, theft, forgery, duplication, guessing, or mimicry [46]. Ideally such a system would allow for conditions where users are under duress and enable long-term auditing. This will reduce vulnerability to authentication that is somehow unintentional [11, 41]).

Assumptions.

The system, including communications between remote sensors, must be measurably secure and trustable by both parties involved in the authentication. It should also tolerate the basic tenets of security including "insider" threats, be they authorized users who have "gone bad," are coerced, or have made mistakes. This can be done, for example, by authenticating several people and requiring consensus among a majority of those people for an action to take place. We assume that electronic communications are digitally signed using means that are not easily forgeable. The implementation of these assumptions are beyond the scope of this paper, however. Finally, we take the original establishment of "identity" as axiomatic: we are concerned only in connecting a person with their attempt to authenticate and not whether the means to do so also establish whether or not the person is the same person that is listed on a birth certificate.

4. PREMISE OF SOLUTIONS

We now discuss the premise of some possible solutions that fit our goals and why a number of alternatives do not.

4.1 Multi-Dimensional, Measurable Inputs

Collectively, rich authentication technologies, including biometrics, facial, environmental, voice recognition, GPS, etc..., are readily available, and can provide measures of identity that approach that of the physical world. Many smart-phones incorporate cameras, GPS receivers, accelerometers, and rate gyroscopes that can identify location [23,30], and evaluate gestures [36], daily movements [43], and writing style [35]. Additional, trivially available sensors such as galvanic skin response and eye tracking [27] enable continuous real-time user authentication. Early steps in rich authentication have been promising ("Cell phones show human movement predictable 93% of the time." [43]).

Of all of these techniques, however, the only individual class of techniques resistent to all of these threats, in principle, are biometrics. And, not only are biometrics capable of being resistent to these threats, but they are *measurably* resistent. That is to say, while there is no way to accurately predict how likely it is, even in the most crucial cases, that someone will forget their password or lose their token [2, 4]. In contrast, biometrics offer distinct advantages over traditional authentication schemes because we can measurably predict the rate of false positives and false negatives based on the type of biometric used [1, 13, 24, 38]. Additionally, biometrics do not have interdependencies such as when one input to authentication depends on another.

But biometrics applied incorrectly can still be seriously flawed. For example, in a situation without a measurably secure biometric reader that the person authenticating trusts, the person authenticating risks that their biometric may be captured and replayed in the future. And, in a situation without a measurably secure biometric reader that the person or system being authenticated to trusts, the person being authenticated to risks being the target of a replay attack. Thus, both sides must have a probably trusted means of reading biometrics and a trusted path between the sensor and the person validating the sensor results, such as the *clean slate* solution referred to earlier [40]. And moreover,

biometrics still result in a degree of confidence, not an absolute certainty.

Thus, we return to our previous assertion that computers lack the intuition that humans can benefit from. This does not mean that computers should not be part of the authentication equation. They are very effective in data correlation and tracking and should be used where they are strong. We believe that their role should focus on providing information, not deciding what to do with that information [22]. Thus, a computer can provide this degree of confidence to a human, but ultimately, a well-trained human is best able to make the decision. For this reason, audio and video should be also communicated as a means of providing as much as possible of an in-person, human-to-human authentication as possible [34]. For example, authentication of an email messages can be practically validated by encoding that message with a continuous video of a person typing the email (and indicating keystroke cadence) and then entering their thumbprint and retina scan. In the process, this communicates several types of biometrics all containing inputs that with measurable degrees of confidence, and, in the process that is combined with human visual and audio cues. While one may comment that these video streams could be faked, we reiterate that our assumption is that trust in the hardware is accomplished through previously-presented means [40].

Biometrics and video—even several biometrics fused together to make masquerade harder—are not sufficient to verify intent, however. For this reason, a "secret" of some kind must also be used so that intent to access can be distinguished from accessing under duress [11, 41]). Such a secret could be a password, but a password to disambiguate intent from duress need not be one that can withstand months of brute force attempts to guess the password [12]. That is not the objective here. The *biometrics* are the system used to provide measurable confidence of identity. The secret or password must simply enable communication of the users intent in a way that would be unlikely to be easily guessed [44]. In theory, such a secret could even be the knowledge of which finger to use to authenticate, with one finger indicating legitimate intent and another indicating duress. The space of finger combinations is probably too small to reasonably prevent guessing, however. But even a simple three-digit code[1] is unlikely to be guessed in the time that intent is communicated. Such a code need not have combinations of digits, punctuation, and upper and lower case letters, and need not be changed ever six months to protect against brute force attacks on weak hashing algorithms [42]. Nevertheless, simple secrets can be vulnerable to unauthorized disclosure which is why they must be used in conjunction with other techniques and why appropriate safeguards are still important.

"Usability" is not central to our theme but it is noteworthy. Currently, we have defaulted to a rather limited condition where we mistake burdensome security for good security: the assumption is that the more burdensome security measures are on authorized users, the more secure we are against unauthorized users [7]. We're finding that this is not the case [21], and is often counterproductive [16, 42]. However, our proposed solution not only provides means to measure how likely someone is who they claim to be, but does so with less burden than existing means. Someone wishing to authenticate with our scheme could literally walk into a room naked, carrying nothing, remembering virtually nothing, and can still authenticate.

Somewhat related to usability is the concept of "break-the-glass" situations such as extremely time-sensitive emergencies when full authentication is not possible and the safety condition allows and requires the full authentication to be bypassed. In such situations, the techniques that we propose in this paper can still be used for *post hoc* auditing and analysis even if they are allowed to be bypassed for the actual authentication [40]. That is, whatever data can be collected can be used to show after the fact if the person who "broke" the glass should have been allowed access to the system.

Finally, we note that we must measure that trust *continuously* and not just at the start of a session. For example, actions, such as the act of sending a message, are not limited to the actual process of pressing the button that sends the message, but also include the process of writing the message. *Continuous* or *dynamic* authentication is not a new concept [29], but is a particularly essential one to the paradigm that we propose, which is based on risk and confidence measures that can and often do change over time. Continuous authentication is a critical component to any resilient solution. Such approaches move beyond traditional passwords and hardware tokens, which capture only an single moment of trust in a particular context. Continuous authentication runs in the background, authenticating the user regularly (e.g., every keystroke, movement) and validating the fact that the user is in the room with a high degree of confidence. This is a simple and very effective part of most face-to-face communications. Continuous authentication may even provide a range of responses, depending on confidence (e.g., not just "allow" or "deny") [39]. Such a gradated response is an intrinsic and normal part of most face-to-face communications yet rather limited in the digital world. It is certainly not unreasonable for a bank teller to ask for more identification when a customer's interactions seem suspicious or they request to move larger sums of money. We believe such continuous and responsive authentication should be an integral part of the digital world.

Simple biometric validation has been exhaustively researched and limitations of such methods are well-known. Biometrics alone are limited in their ability to provide robust authentication. Nevertheless, biometrics are an essential element of any comprehensive authentication scheme. The continuous, real-time fusion of biometrics to validate identity and a very simple secret to validate intent can be used to improve authentication and reduce ambiguity while also improving the probability that the authorized user is who they claim to be, and intends to access a resource not just at the start of an authenticated session but throughout it.

It is important to note that any attempt to provide robust authentication is likely to erode privacy, due to fundamentally contrasting goals. We consider this balance intuitively every time we choose to physically attend a meeting. In addition to verbally sharing our thoughts, physical attendance demands that we reveal much additional information about us. We are for example, revealing what we look and sound like, our location, mannerisms, and more. In the case of a face-to-face meeting, we often unconsciously make the decision that the value of attendance is worth the compromise in privacy.

[1] Auto-destruct:
`http://en.memory-alpha.org/wiki/Auto-destruct`

4.2 State and Behavior Metrics

In addition to those metrics that come from a specific, individual biological trait, we can build or erode confidence based on combined state and behavioral metrics. Adding combined state and behavioral metrics to a more traditional biometric results in the following notional taxonomy:

1. Something you are
   (a) Temporal issues*
   (b) Spatial-temporal consistency*
      i. Time history of biometrics compared to capabilities of the human body.
   (c) Biometrics (Instantaneous capture/assessment where each of these is a function of each keystroke or other system input)
      i. Fingerprints while typing each keystroke
      ii. Keystroke dynamics*
      iii. Facial/iris/retina recognition
      iv. Instantaneous current location (e.g., via GPS, camera)
      v. Voice/grammar/idiom recognition*
      vi. Gait
      vii. Body dynamics (e.g., how a smartphone is held or moved)
2. Something you know
   (a) Password/passphrase
   (b) Analog combination
3. Something you have
   (a) Physical key
   (b) One-time pad
   (c) Hardware two-factor authentication token

Building on the canonical face-to-face interaction we now discuss several possible combined state and behavioral metrics (marked with a * in the notional taxonomy above).

Temporal issues.

Some authentication techniques are continuous and unbroken. An example of this is a video stream. Others are ongoing but are discrete events. An example of this is keystrokes. Yet other techniques reflect individual points in time and are completely discrete events. An example of this is traditional password entry. The simple fact that time has elapsed since the last face-to-face encounter serves to erode authentication confidence. Additionally, each of these other categories can also reflect the erosion of trust since the last face-to-face encounter. This can be done as a function of time (your login is only good for 10 minutes) or as a function of activity (e.g., idle terminal auto logout). Ideally, more gradual erosion could offer less critical functionality or require re-authentication when critical tasks were requested.

Spatial-temporal consistency.

Identity is compromised if a person is perceived to appear in two places at once. For example, a near-simultaneous appearance in geographically separated locations can indicate that an attack is underway. An individual logging in from to geographically distant locations moments apart could indicate the compromise of a shared secret, such as from as a keystroke logger on the first machine that was used to obtain a password to enable login from the second, geographically-distant machine. Again, used on its own, "state" consistency is likely to be insufficient. Fusion is key among "state" input data as well. Knowing a signal is coming from a home need not mean that the person attempting to authenticate is the person expected—it could be a spouse, a child, a burglar, etc... The more the information provided uniquely identifies the person in question, clearly the more valuable the information [50]. And, additionally, not all "state" metrics are necessarily even useful, either. For example, GPS is likely to be much more meaningful than "IP geolocation," which can be trivially thwarted with VPNs, proxies, etc... [33].

Starting from the face-to-face encounter, a spatial probability "basket" can be developed as a function of time based on the equations of state for human mobility. An individual appearing outside this spatial probability basket would erode authentication confidence.

Note that the "consistency" we refer to here is consistency as constrained by the laws of physics and not predictability, for example, if a user is simply trying to authenticate from a place that they do not usually authenticate from, much as is used with credit card fraud detection. Such a measure may be valuable in some cases but it is anathema to the system that we propose that needs to provide measurable certainty of identity. Such a system cannot reduce certainty just because of unusual circumstances because the system itself may need to be usable under unusual circumstances.

Biometrics.

Additionally, a variety of biometrics are frequently not considered in practice as things like fingerprint readers have but also have a measurable and scientific basis:

Use of grammar and idiom. Individuals use language differently – speech patterns represent something analogous to a psychological biometric. Grammar checkers analyze text based on accepted rules of proper usage. These grammar rules, however, allow significant freedom. Within these rules, individuals develop specific writing styles that are recognizable. For example, when reading the words "Speak, friend, and enter." [47], should one interpret this to mean "Say the word 'friend' and then enter" or "Friend! Say something and then go inside"? Voice-recognition software exploits this individually specific style to improve speech recognition. In the same way, individual writing style can be used to increase authentication confidence.

Keystroke dynamics. Beyond a lingual analysis, keystroke dynamics (and other inputs to computers) have been shown to be a useful tool for authentication [26]. Such a system provides greater confidence with more keystrokes. Enabling such monitoring is yet another real-time continuous metric that can be used to increase or erode trust in a users identity.

As mentioned above, these metrics are fabulously invasive. Thus, they may only be desirable for high assurance, critical systems. Nevertheless, an evaluation similar to that by Bonneau, et al. [10] reveals that they would provide dramatic improvements in security with only limited compromises in usability and deployability.

Our goals combined with the current limitations of science and technology lead to a set of principles that we assert that systems must adhere to in order to capture the confidence measures of elements that make the current *de facto* standard of in-person, human-to-human function well.

5. PRINCIPLES OF AUTHENTICATION

We posit the *Principles of Authentication* that describe means of validating the amount of trust that one can place in a process of authentication: Recall that we treat *identity* as axiomatic. Identity could be a person, organization, property, set of properties, etc... Given that, the following principles speak to authenticating that identity.

0. Identity should be verified as long and as frequently as access to a resource is permitted. If access is ongoing then identity verification should be continuous.

1. Authentication is about validating whether or not someone is who they claim to be, and about determining whether that person intends to authenticate and is not, for example, being coerced.

2. In person, human-to-human authentication is the *de facto* standard in life. When this is not possible and computers must be involved, then computers should provide measures of confidence (or lack thereof) to humans. Those humans should ultimately make authentication *decisions*, not computers.

3. Authentication should be trivial for the person legitimately authenticating but hard for an adversary to defeat, much like verifying a key vs. guessing a key in public key cryptography.

Authentication that scales builds on these principles through bootstrapping. Using these techniques, we assert that *humans* should judge things based on the confidence level a computer provides. Moreover, instead of a single sign-in event enabling access until the user logs out, rich authentication intelligently fuses sensor data with predictable human behavior and limitations to enable probabilistic (and difficult to mimic) confidence that the specific user is at the machine. For example, conceivably, every keystroke [25] and mouse motion can be automatically and transparently signed indicating the confidence that the machine input was provided by the specific individual. If confidence is eroded due to user inconsistency (such as injury or sickness), confidence can be restored by requiring alternative (and potentially more intrusive) means such as DNA analysis of blood samples (as such technologies become widely available).

We reiterate that one of the reasons that humans are able to do this is *context*. We observe that in the physical world, *physics* is relevant. Humans can take this into account. Computers can too, but how much should they weigh it? For this reason, we believe authentication should employ additional factors that enable confidence about identity to be increased or decreased. Additional factors could include: *where you are* (e.g., via GPS) [14]—because human motion is governed by the laws of physics; and *time*—it is physically impossible for a human to be in two places at once, and so if Person A was in California at 10:00 AM on Tuesday, they won't be in Beijing one hour later.

Moreover, we observe that in the physical world, "authentication" (recognizing someone) draws heavily on intuition. In fact, humans perform this function naturally and trivially. In contrast, machines cannot understand the subtlety of authentication. For example, they cannot easily interpret the meaning of being "Facebook friends" with someone. Again, in contrast, humans can recognize gait, posture, and other forms of "body language," even via videoconferencing.

Alternatively, suppose that there exists background noise in a phone call from Disneyland. Humans recognize such sounds and intuitively and unconsciously check to ensure that this background information is consistent with the authentication "picture." For example, if an individual in a phone call claims to be delayed at the airport, sounds that are inconsistent with an airport (i.e., sounds from Disneyland) would serve to erode confidence in this authentication picture. Other participants in the phone call may choose to increase confidence in identity by pointing out this inconsistency. A simple and unobtrusive inquiry ("If you are at the airport, why do I hear Disneyland sounds in the background?") followed by a credible response ("Oh, I'm walking past the Disney store in the airport lobby right now.") would serve to increase confidence. Therefore, our idea is to present rich information to other human users to enable detection of inconsistencies in authentication.

Note that the "consistency check" provided by such background noise is anathema to computing systems. For example, Google Voice transcribes phone messages into text which is then e-mailed to the recipient. Background noise in a phone message complicates the speech recognition process and is unnecessary for the transcription to text. Therefore, the computing system regards such background noise as a complicating factor to be filtered out. As a result, the transcribed message loses a rich source of authentication information that could provide a consistency check.

An illustration of how a human might use this information is shown in Figure 1. In that figure, "From Sean with 80% confidence" denotes that an e-mail has been received 'from Sean' just like any other e-mail application. The "80% confidence" notation is the summary metadata from all of the authentication confidence metrics described in this paper. The recipient of the e-mail may choose to open the e-mail message without any further consideration if 80% confidence is sufficient to meet their needs.

In the event that 80% confidence "seems a little low" based on previous context (such as e-mails received from this sender in the past or the criticality of the information in the e-mail) the recipient may desire increased confidence that this e-mail is, indeed, from Sean. In this situation, the user can click on the confidence and be taken to a graphical representation of the metadata used to produce the "80% confidence" metric. In Figure 1, this notional graphical representation summarizes the metrics leading to "80% confidence"—"'biometrics" and "equations of state." Clicking on "biometrics" takes the user to a comparison of Sean's biometrics from this e-mail with biometrics from previous e-mails and supplemental biometric updates from Sean. In this notional example, there is a 98% match between current and previous biometrics which the recipient could consider good enough for confidence in this message.

The "equations of state" metric, however, is only at 50% due to the fact that Sean has been moving at a high rate of speed since previous e-mails and updates were received from Sean. "Sean's estimated location" is based on the maximum reasonable speed (e.g., on an airplane) at which Sean could be traveling for the amount of time since the last update—hence the circle shown on the map would be increasing in diameter with time. Since the graphical representation of "Sean's current location" is at the edge of (but, nevertheless, within) this circle, the recipient may be led to conclude, correctly, that Sean has been traveling.

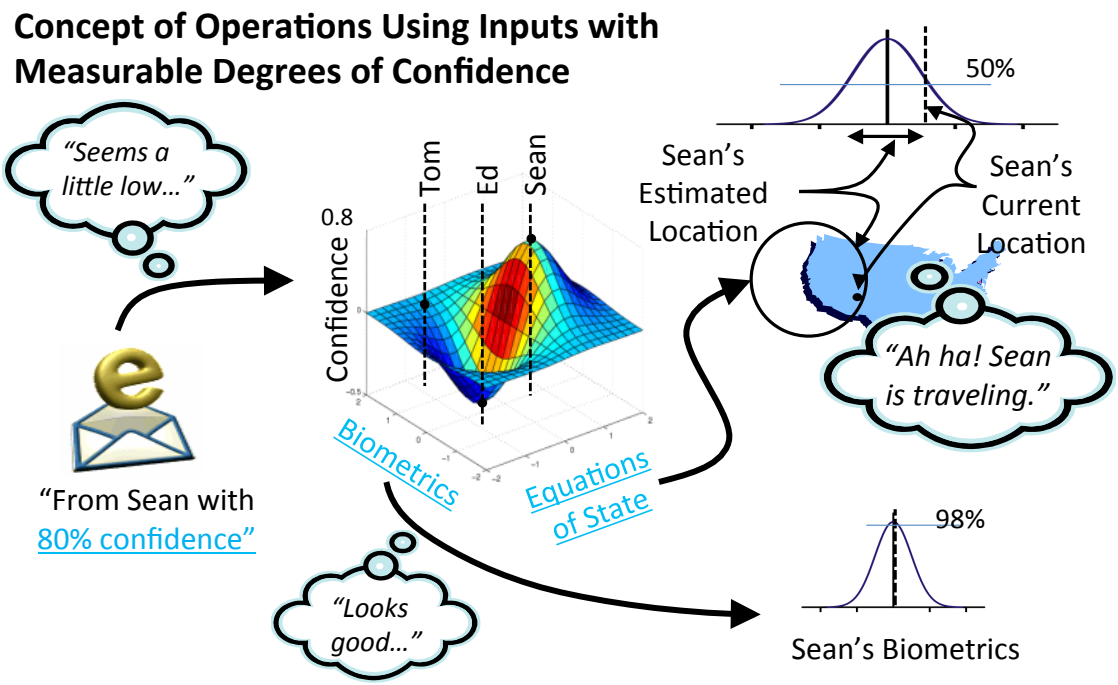

Figure 1: An example of a concept of operations for authentication using input factors that provide meaningful confidence measures.

If even greater confidence is needed (again, due to the criticality of the information in the e-mail) the recipient may choose to call Sean on the cell phone and verbally verify that he is somewhere in the vicinity of Albuquerque, New Mexico. The recipient may even choose to further corroborate this cell phone conversation with Sean by calling some of Sean's associates in Albuquerque.

6. EXAMPLES

In this section, we consider three alternative applications for continuous or dynamic authentication.

Human-Machine Authentication.

Continuous authentication develops a probability of user presence that moves up and down depending on the situation and as new sources of data are added. As a user walks into the same room with the "authenticator device" (a computer, tablet, smart phone, etc.), the authenticator device may use a facial recognition algorithm to develop a probability of the user being in the room of 0.7. As the user picks up the authenticator device, the device may use accelerometers and rate gyros to recognize gestures invoked by the user to increase this probability to 0.85. As the user moves around, the authentication device may use the same devices to increase the probability that the user is in the room to 0.93 over time. After two hours, continued observation of the user's biometrics may increase the probability that the user is in the room to 0.97. When the user puts on the iris/retina recognition/eye-tracking glasses, the probability of the user's presence goes to 0.993. During all of this, the user might be dictating to a computer. Based on the fact that the human does not use any "duress words" during that time, the probability of duress is low. After the user sends a couple of emails, observation of the user's typing rhythms provides a probability that those emails come from the user and that the user is sanguine goes to 0.99995.

Human-Human Authentication.

Humans are far more capable authenticators than machines. However, there are also limitations to a human's ability to perform authentication. For example, it has been shown that humans can only maintain stable relationships with roughly 150 people [15]. Since authentication may require going outside that bound, it is necessary to develop means for providing that ability. Modern social networking sites are an example of something that gives a means for one person to reference beyond the 150 people that they know personally by using people within their 150-person network to validate and vouch for people outside that 150 people.

Additionally, the sensor-fusion assessment described above can provide useful information to aid human authentication. For many interactions, the authentication assessment prepared for human-machine authentication may be more than sufficient. In addition, the human-machine authentication can keep the user from making stupid mistakes.

For critical interactions in a high-threat environments, however, raw information (such as the video feed of the user walking into the room, etc.) should be provided with high integrity to the decision-maker. In such situations, inconsistencies between the human-machine authentication result and the raw information provided to the decision-maker may serve to erode confidence below that developed through human-machine authentication. In the same way a heads up display helps pilots track aircraft status and select targets we envision an environment where a machine may help the

humans process the data but in the end they select the target and they verify the authentication.

Machine-Human Authentication.

We note in passing that there are examples of where a machine may need to authenticate to a human, for example a remote sensor in a hostile environment. While this is a related issue to what we discuss in this paper because it also involves *origin integrity*, the topic is mostly out of the scope of this paper because we believe that the solutions relate less to verifying identity and more to the integrity of the data and/or sensor. Thus, we feel machine-human authentication relates more to a combination of data provenance, our previously-described approach on clean slate designs [40], and, in some cases, procedures [37] similar to zero-knowledge protocols [18].

7. SUMMARY

Online activities can approach the level of clarity, certainty and intuitiveness as activities in the physical world. Physical world metaphors drive the entire user experience. However, the misapplication of some of these metaphors physical metaphors—e.g., resemblance as opposed to mere consistency—can create anxiety and a lack of clarity for users about online actions. Moreover, a mismatch in goals—e.g., preventing attack of a captured set of password hashes, rather than validating user identity—lead to solutions that are inappropriate in some situations, and certainly in critical environments. Our principles of authentication are a solution to this mismatch. By properly ensuring consistency between two worlds and appropriately managing the role of humans vs. the role of computers, the "membrane" between the physical and online world effectively disappears.

Acknowledgements

This research was supported in part by the Director, Office of Computational and Technology Research, Division of Mathematical, Information, and Computational Sciences of the U.S. Department of Energy, under contract number DE-AC02-05CH11231. It was also supported in part by the National Science Foundation and the GENI Project Office under Grant Number CNS-0940805, and by the National Science Foundation under Grant Numbers CCF-1018871, CCF-1049738, and CNS-1312573.

Some of this work was completed while Ed Talbot was a Distinguished Member of Technical Staff at Sandia National Laboratories. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

We are grateful to the attendees of NSPW 2013 in Banff for the lively discussion and valuable feedback that helped to improve this paper. We are especially grateful to Prof. Matt Bishop of UC Davis for his particular insights into some of the subtle but extremely important nuances of the problems and solutions that we've tried to address in this paper.

Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect those of any of the employers or sponsors of this work.

8. REFERENCES

[1] P. S. Aleksic and A. K. Katsaggelos. Audio-Visual Biometrics. *Proceedings of the IEEE*, 94(11):2025–2044, Nov. 2006.

[2] M. Ambinder. Why Clinton's Losing the Nuclear Biscuit Was Really, Really Bad. The Atlantic, Oct. 22 2010.

[3] R. J. Anderson. *Security Engineering: A Guide to Building Dependable Distributed Systems, Second Edition*, chapter 13, Nuclear Command and Control. Wiley Publishing, second edition, 2008.

[4] BBC News. Clinton drops nuclear football. `http://news.bbc.co.uk/2/hi/americas/328442.stm`, Apr. 26 1999.

[5] S. Bellovin and G. McGraw. Silver Bullet Show 081 - An Interview with Steve Bellovin. `http://www.cigital.com/silver-bullet/show-081/`, Dec. 26 2012.

[6] M. Bishop. *Computer Security: Art and Science*. Addison-Wesley Professional, Boston, MA, 2003.

[7] M. Bishop. Chapter 1: Psychological Acceptability Revisited. In L. Cranor, editor, *Security and Usability: Designing Secure Systems That People Can Use*. O'Reilly and Associates, Inc., 2005.

[8] M. Bishop, S. Engle, S. Peisert, S. Whalen, and C. Gates. We Have Met the Enemy and He is Us. In *Proceedings of the 2008 New Security Paradigms Workshop (NSPW)*, pages 1–12, 2008.

[9] J. Bonneau. The Science of Guessing: Analyzing an Anonymized Corpus of 70 Million Passwords. In *Proceedings of the 33rd IEEE Symposium on Security and Privacy*, pages 538–552, May 2012.

[10] J. Bonneau, C. Herley, P. C. van Oorschot, and F. Stajano. The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes. In *Proceedings of the 33rd IEEE Symposium on Security and Privacy*, pages 553–567, May 2012.

[11] D. Chaum. Security Without Identification: Transaction Systems to Make Big Brother Obsolete. *Communications of the ACM*, 28(10):1030–1044, 1985.

[12] J. Clark and U. Hengartner. Panic Passwords: Authenticating Under Duress. *Proc. of the 3rd USENIX Workshop on Hot Topics in Computer Security*, 2008.

[13] J. Daugman. How Iris Recognition Works. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):21–30, Jan. 2004.

[14] D. E. Denning and P. F. MacDoran. Location-Based Authentication: Grounding Cyberspace for Better Security. *Computer Fraud & Security*, 1996(2):12–16, 1996.

[15] R. I. M. Dunbar. Neocortex Size as a Constraint on Group Size in Primates. *Journal of Human Evolution*, 22(6):469–493, 1992.

[16] S. Egelman, D. Molnar, N. Christin, A. Acquisti, C. Herley, and S. Krishnamurthi. Please Continue to Hold: An Empirical Study on User Tolerance of Security Delays. In *Proc. of the Ninth Workshop on the Economics of Information Security (WEIS)*, 2010.

[17] S. Egelman, A. Sotirakopoulos, I. Muslukhov, K. Beznosov, and C. Herley. Does My Password go up

to Eleven? The Impact of Password Meters on Password Selection. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, 2013.

[18] S. Goldwasser, S. Micali, and C. Rackoff. The Knowledge Complexity of Interactive Proof-Systems. In *Proc. of the 17th Annual ACM Symposium on Theory of Computing*, 1985.

[19] D. Goodin. Anatomy of a Hack: How Crackers Ransack Passwords Like "qeadzcwrsfxv1331". *Ars technica*, May 27 2013.

[20] E. Grosse and M. Upadhyay. Authentication at Scale. *IEEE Security & Privacy*, 11(1):15–22, Jan.-Feb. 2013.

[21] C. Herley. So Long, and No Thanks for the Externalities: The Rational Rejection of Security Advice by Users. In *Proc. of the 2009 New Security Paradigms Workshop (NSPW)*, 2009.

[22] G. Holland and L. Burton et al. Dreadnought (Star Trek: Voyager). Paramount, February 1996.

[23] F. Hsu, H. Chen, and S. Machiraju. WebCallerID: Leveraging Cellular Networks for Web Authentication. *Journal of Computer Security*, 19(5):869–893, 2011.

[24] A. K. Jain, A. Ross, and S. Pankanti. Biometrics: a Tool for Information Security. *IEEE Trans. Inf. Forensics and Security*, 1(2):125–143, 2006.

[25] K. S. Killourhy and R. A. Maxion. Comparing Anomaly-Detection Algorithms for Keystroke Dynamics. In *IEEE/IFIP International Conference on Dependable Systems & Networks (DSN)*, 2009.

[26] K. S. Killourhy and R. A. Maxion. Free vs. Transcribed Text for Keystroke-Dynamics Evaluations. In *Proceedings of the 2012 Workshop on Learning from Authoritative Security Experiment Results (LASER)*. ACM, 2012.

[27] M. Kumar, T. Garfinkel, D. Boneh, and T. Winograd. Reducing Shoulder-Surfing by Using Gaze-Based Password Entry. In *Proc. of the 3rd Symposium on Usable Privacy and Security*, 2007.

[28] A. J. Lee and M. Winslett. Enforcing Safety and Consistency Constraints in Policy-Based Authorization Systems. *ACM Transactions on Information and System Security (TISSEC)*, 12(2):1–33, 2008.

[29] J. Leggett, G. Williams, M. Usnick, and M. Longnecker. Dynamic Identity Verification via Keystroke Characteristics. *International Journal of Man-Machine Studies*, 35(6):859–870, 1991.

[30] E. Lin, J. Aycock, and M. Mannan. Lightweight Client-Side Methods for Detecting Email Forgery. In *Proc. of the 13th International Workshop on Information Security Applications (WISA)*, 2012.

[31] T. F. Lunt and R. Jagannathan. A Prototype Real-Time Intrusion-Detection Expert System (IDES). In *Proc. of the IEEE Symposium on Security and Privacy*, pages 59–66, 1988.

[32] A. Mehrabian. *Silent Messages*. Wadsworth, 1971.

[33] J. A. Muir and P. C. van Oorschot. Internet Geolocation: Evasion and Counterevasion. *ACM Computing Surveys (CSUR)*, 42(1), 2009.

[34] D. Nali, P. C. van Oorschot, and A. Adler. Videoticket: Detecting Identity Fraud Attempts via Audiovisual Certificates and Signatures. In *Proc. of the New Security Paradigms Workshop*, 2007.

[35] A. Narayanan, H. Paskov, N. Z. Gong, J. Bethencourt, E. Stefanov, E. C. R. Shin, and D. Song. On the Feasibility of Internet-Scale Author Identification. In *Proc. of the 33rd IEEE Symposium on Security and Privacy*, May 2012.

[36] Y. Niu and H. Chen. Gesture Authentication with Touch Input for Mobile Devices. In *Proc. of the 3rd Intl. Conf. on Security and Privacy in Mobile Information and Communication Systems*, 2011.

[37] Office of the Director of National Intelligence. United States Intelligence Community Information Sharing Strategy. `http://www.dni.gov/reports/IC_Information_Sharing_Strategy.pdf`, Feb. 22 2008.

[38] L. O'Gorman. Comparing Passwords, Tokens, and Biometrics for User Authentication. *Proceedings of the IEEE*, 91(12):2021–2040, Dec. 2003.

[39] S. Peisert and M. Bishop. Dynamic, Flexible, and Optimistic Access Control. Technical Report CSE-2013-76, University of California at Davis, March 2013.

[40] S. Peisert, E. Talbot, and M. Bishop. Turtles All The Way Down: A Clean-Slate, Ground-Up, First-Principles Approach to Secure Systems. In *Proc. of the New Security Paradigms Workshop*, pages 15–26, 2012.

[41] G. Roddenberry and G. L. Coon et al. Bread and Circuses (Star Trek: The Original Series). NBC, March 1968.

[42] A. Singer and W. Anderson. Rethinking Password Policies. *;login:*, 38(4):14–18, August 2013.

[43] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási. Limits of Predictability in Human Mobility. *Science*, 327(5968):1018–1021, 2010.

[44] S. Spielberg and G. Lucas, *et al.* Indiana Jones and the Last Crusade. Lucasfilm Ltd., 1989.

[45] S. D. Spray and J. A. Cooper. The Unique Signal Concept for Detonation Safety in Nuclear Weapons. Technical Report SAND91-1269, Sandia National Labs., Albuquerque, NM (United States), June 1993.

[46] E. B. Talbot, D. Frincke, and M. Bishop. Demythifying Cybersecurity. *IEEE Security & Privacy*, 8(3):56–59, May/June 2010.

[47] J. R. R. Tolkien. *The Fellowship of the Ring*. George Allen & Unwin, 1954.

[48] A. M. Turing. Computing Machinery and Intelligence. *Mind*, 59(236):433–460, 1950.

[49] P. C. van Oorschot. Message Authentication by Integrity with Public Corroboration. In *Proc. of the New Security Paradigms Workshop*, pages 57–63, 2005.

[50] P. C. van Oorschot and S. Stubblebine. Countering Identity Theft Through Digital Uniqueness, Location Cross-Checking, and Funneling. In *Proceedings of the 9th International Conference on Financial Cryptography and Data Security*, pages 31–43, 2005.