A Survey on User-Device Authentication on Emerging HCI Interfaces

Abstract—As user demand and cost benefits of natural user interface technologies are hastening their wide adoption, computing devices come equipped with these natural interaction interfaces are becoming ubiquitous. Consequently, authentication mechanisms on them are becoming an essential security component to allow wider range of applications to be deployed on the devices including privacy and security sensitive ones.

This paper presents a comprehensive review of existing user device authentication approaches that are applicable on new emerging interfaces and identifies their security and usability questions that shall be pursued in future work. In addition, a set of evaluation metrics to assess security and usability of these approaches is derived and a comparative evaluation of different authentication approaches is performed. Finally, future research direction in this topic is discussed. The emerging interfaces include touch-sensitive display, 2-d and 3-d camera, voice interfaces, eye tracking platforms, and brain-computer interaction interfaces.

Index Terms—IEEEtran, journal, \textit{\LaTeX}, paper, template.

I. INTRODUCTION

Recent advances in natural user interface technology have remarkably transformed the way we interact with computing devices and the form factors of the computing devices themselves. We are now surrounded by many new categories of computing devices including tablets, tabletops, wearable and augmented reality devices, where we could interact with them more transparently and seamlessly using new forms of natural interfaces. These include, but are not limited to, touch gestures and operated on these interfaces and devices.

However, without user device authentication mechanisms in place, security and privacy of users associated with sensitive information stored in or accessible from the device will be at risk of being compromised and violated. In addition, types of applications users choose to deploy on these devices will be of the limit. In 2012, Symantec has performed an experiment to study behaviors of strangers who picked up unlocked phones in several cities of the US and Canada [1]. Interestingly, they found that 96% of the devices were accessed by the finders whereas only 50% of the finders contacted the owner and provided contact information. More over, the recent study by Egelamn et al. [2] has indicated that at least over one third of the users stored their sensible data, e.g., SSN, last four digits of SSN, bank account, credit/debit card number, DOB, email password, and home address in their email accounts, which are most likely accessible from their smartphones. And the fact that this information can be found in their emails was a surprise to them. In addition, the same study also reported that users who did not locked their smartphones stored significantly less sensitive information on their emails than those who locked their devices. This might imply that use users who do not have locking mechanism in place may deter to install or use some sensitive applications, e.g., mobile payment.

On the other hand, the age-old text password mechanism is facing new usability challenges, in addition to its known security and usability issues, when deploying on new forms of computing devices that are not equipped with a physical keyboard. This in turn results in users ignoring this mechanism. According to the study on user locking behavior on Android’s touch-sensitive smartphones, only a small percentage of users (14%) chose to deploy this text password mechanism whereas a much larger percentage of the users (51%) chose Android’s Pattern Lock as an unlocking mechanism [3]. More interestingly, there is currently no user authentication mechanisms implemented on off-the-shelf Google augmented reality headset namely Google Glass, and Samsung smartwatch namely Galaxy Gear.

As such, there is a window of opportunity for ones to develop user-device authentication mechanisms on emerging interfaces which can then be operated on new computing devices and platforms. And with the industrial effort to standardize web authentication protocol driven by FIDO alliance, not only such a mechanism can be used for local device authentication, it also be used to initiate the online or web authentication as well [4]. In response, a number of authentication alternatives that are applicable on these emerging interfaces have been proposed and investigated. In addition, several variants and incremental development of these approaches have been suggested and their related studies have been conducted. Nevertheless, the studies related to these approaches often preformed and analyzed independently and diversely by experts in various communities. As a result, the true merit of the approaches can not be justified since some essential aspects of usability and security were missing from their analysis [5].

The focus of this paper is to provide a systematic review of existing authentication approaches that are applicable on different types of emerging interfaces and to identify key usability and security challenges of these approaches. In addition, a comprehensive set of criteria to assess security and usability of user-device authentication mechanisms is presented along with the comparative evaluation of these approaches. This information could help researchers to refine the authentication mechanisms that are inherently more secure and usable, with respect to the context of use, on new forms of computing devices.

Furthermore, this paper could serve as a starting point to develop a holistic authentication approach on these new devices by taking advantage of available interfaces to enhance security.
TABLE I: Summary of authentication approaches

<table>
<thead>
<tr>
<th>Interfaces</th>
<th>Examples of device</th>
<th>Approaches</th>
<th>Related studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Touch interface</td>
<td>Tablet, Smartphone, Smartwatch, Multitouch coffee table, and Large interactive display</td>
<td>-Textual password on a virtual keyboard</td>
<td>[6], [7]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Android screen-lock pattern</td>
<td>[8]–[10]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Microsoft picture</td>
<td>[11], [12]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Multi-touch gesture</td>
<td>[13]–[16]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Online signature</td>
<td>[17]–[22]</td>
</tr>
<tr>
<td>2-D camera</td>
<td>Tablet, Smartphone, Smartwatch, and Head-mouthed device, e.g., Google Glass and Oculus Rift</td>
<td>-Hand gesture authentication</td>
<td>[23]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Face recognition</td>
<td>[24]–[26]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Fingerprint recognition</td>
<td>[27], [28]</td>
</tr>
<tr>
<td>3-D camera</td>
<td>Body motion control, e.g., Kinect and Hand motion control, e.g., Leap</td>
<td>-In-air signature</td>
<td>[29], [30]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-SignWave</td>
<td>[31], [32]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Body gesture</td>
<td>[33], [34]</td>
</tr>
<tr>
<td>Voice</td>
<td>In-vehicle controller</td>
<td>-Voice authentication</td>
<td>[35]–[41]</td>
</tr>
<tr>
<td></td>
<td>Smartphone and Head-mouthed device</td>
<td>-Whisper authentication</td>
<td>[42]</td>
</tr>
<tr>
<td>Eye tracking</td>
<td>Smartphone and in-vehicle controller</td>
<td>-Gaze authentication</td>
<td>[43]–[47]</td>
</tr>
<tr>
<td>Brain interface</td>
<td>EEG headset</td>
<td>-Brain wave authentication</td>
<td>[48]–[50]</td>
</tr>
</tbody>
</table>

provided to users without trading-off usability. That is, as these new devices are equipped with various types of user interfaces, there is a larger amount of user interaction’s trait that could be silently observed and make use by the authentication mechanism to ensure the identity of users. The summary of authentication approaches on emerging HCI interfaces reviewed by this paper is presented in Table II.

A. Background

Computing devices are traditionally equipped with a physical keyboard. Consequently, textual password which can be input by typing on the keyboard has been used as a de facto user authentication mechanism, despite its known security and usability problems (a memorability issue in particular). However, the change of user interaction interfaces and the usage pattern of these new devices have brought new security threats and usability issues to this conventional authentication mechanism, namely text password. First, these devices are often used in social context and in public space where users are surrounded by handheld recording devices and public surveillance. In such situations, the assumption that users can enter a password or other shared-secret credentials to the devices privately and securely no longer holds. In particular, the study by Shukla et al. has reported that 85% of PIN entries could be recovered within ten attempts using only information about dynamics of the hand [51], [52]. Noting that this information can be recorded easily using smart-phones or wearable devices. Secondly, computing devices are no longer built around a physical keyboard as a main user interaction interface but they employ other natural user interfaces (NUI). As a result, there is an inevitable end to the era of user inputting text password on a physical keyboard.

Moreover, since these new interfaces are not designed for text oriented applications, designing usable mechanism to input text password is proven not trivial from both user and machine perspective. In particular, research has shown that entering text password on a virtual keyboard is much slower and harder [53] than on a physical keyboard due to fat finger problem [54], user distraction [55], and inaccurate touch input precision [56], [57]. Consequently, most users do not choose text passwords as an unlocking mechanism on their multi-touch devices [11].

Therefore, a number of authentication alternatives on these interfaces have been proposed and investigated. In addition, many techniques that take advantages of these emerging interfaces to provide intrinsic security to users have been developed. That is, user interaction on these new interfaces provides much higher dimensional data than the one on conventional interfaces. This interaction information can be utilized as additional authentication factors. For example, with a physical keyboard, only press-up and press-down key are detected during user interaction, whereas, with mouse, a time-series of mouse movement and mouse-click event is continuously captured. With multi-touch surface, not only can the system detect multiple of touch points at the time, it can
also collect other contextual information about touch event, for example, touch area, touch pressure, and touch shape. With voice interface, a voice signal, which is one dimensional time-series data of voice amplitudes, is captured with a very high sampling rate. With two dimensional or three dimensional cameras, the output is a time-series of part of the user's body interaction, e.g., whole body and hand gestures. With eye tracker system, a time-series of gaze movement can also be detected. And with the recent commercial brain computer interface (BCI), the sensor system outputs neural signals resulting from user's brain activities which is a time-series of multiple attributes recorded from several EEG sensors attached to user's head surfaces with a sampling rate up to 1 kHz. Using information extracted from user interaction on these emerging interfaces, several studies have shown that biometric traits of users can be derived and it can be used to ensure the identities of users based on who-they-are, in addition to what-they-know.

B. Outline of this paper

The remainder of paper is organized as follows. First, we present a review of authentication approaches on touch interfaces in Section II, and followed by the the ones on 2-D and 3-D sensing cameras in Section III and IV, respectively. Then, those on voice, eye tracking, and brain computer interfaces are presented in section V, VI, and VII. Each of these sections begins by describing the hardware sensor technologies of the user interfaces and the device categories that come equipped with these interfaces. Then, authentication approaches that are applicable on this interface are reviewed. For each authentication alternative approach, security and usability benefits are also analyzed.

Then, security and usability evaluation metrics for user-device authentication mechanism are presented in Section VIII. Lastly, in Section IX, we discuss conclusion remarks and future research direction to develop inherently secure and usable authentication approaches.

II. TOUCH INTERFACE

Touch sensitive display has increasingly become a dominant human computer interaction interface in several device categories including handsets, tablets, large interactive displays, multi-touch coffee tables, and smartwatches. To communicate with devices via touch sensitive display, a user performs touch gestures on the display with one or more fingers from one or two hands. Nevertheless, the type of touch gestures that users can comfortably perform are typically restricted by size of the device’s display. For example, a large multi-touch sensitive display can accommodate more complex touch gestures, e.g., the ones performed by four or more fingers, as opposed to the smaller one where only simple gestures can be performed. In turn, user’s touch input is sensed by an array of capacitive or resistive sensors that is embedded within the display. In the context of a signal, a touch gesture is a time series of touch point sets where each point in the set has attributes such as x-y coordinates, time stamp, and sometimes pressure information.

With respect to authentication mechanisms that have been deployed or proposed, the authentication interaction varies from mimicking keyboard interaction (typing gesture) and mouse interaction (tap, draw, and drag) to multi-touch interaction.

Textual password on a virtual keyboard This is one practical implementation of user authentication mechanisms on touch sensitive display. In this approach, a virtual keyboard or keypad is rendered on the display, as illustrated in Figure 2. A user is then authenticated by typing an alphanumeric password on that virtual keyboard. The system then authenticates users by comparing the typing password and the setting one. In addition, Ben et al. [6] have shown that typing characteristics on soft keyboard (the specific location touched on each key, the drift from finger down to finger up, the force of touch, the area of press) could also be used to differentiate the users. Similarly, the study conducted on 16 subjects has demonstrated that typing gestures of text messages are unique to individuals and they are consistent over time [7]. This implies that soft-keyboard typing behavior can also be used as an additional authentication factor. However, the verification error of the system when only few characters are typed (typical user’s choices of passwords) is very high (32.3% for five keypresses in the first study and ranging from 2% to 50% for a single work in the second study).

Furthermore, while such an approach enjoys the benefit of user’s familiarity, it also inherits several security drawback from the textual password approach including shoulder surfing. In particular, it has been demonstrated that a PIN entry can be automatically recovered from the video recording of the hand dynamic while it is being entered on a smartphone (via google glass, hand-held recording devices, public surveillance, etc.) [51], [52]. While the proposed solution for this is to use randomized keyboard as opposed to the static layout one, a user would have to spend the time as twice as it takes on a normal keyboard and, in addition, the solution does not prevent the attack where the screen content is also visible in the video. Moreover, users often select weak passwords in response to their inability to memorize multiple of complex passwords. In addition, typing on a virtual keyboard has incurred new usability problems. In particular, entering text entry on a virtual keyboard is reportedly slower and harder than on physical keyboard [53]. This is due to fat finger problem [54], user distraction [55], and inaccurate touch input precision [56], [57]. This in turn results in users ignoring this mechanism. According to Bruggen, et al. [3], of the Android’s users who lock their devices, only 22% of them chose to deploy alphanumeric password as their authentication mechanism.

Android screen-lock pattern This graphical password scheme is one variant of draw-a-secret proposed by I. Jermy et al. [8], in 1999. The scheme is regarded as a phone’s unlock mechanism that is more popular than a form of text password [3]. That is 78% and 22% of the users who locked their phones were used a pattern and password as an authentication mechanism, respectively, where 35% of all users did not enable any authentication mechanism. In this scheme, a user is asked to create and memorize an exact dot-connected pattern on 3×3 grid, as shown in Figure 4. The user is then authenticated based purely on knowledge of a drawing pattern shared between the system and the user. As a result, an attacker who observes a user drawing a legitimate pattern
could use that knowledge to gain access to the system. In addition, the recent study by Andriotis et al. [9], in 2013, have shown that by combining physical, i.e., oily residue or smudges, with the knowledge of pattern distribution, 54.54% of the patterns can be fully recovered. To address this problem, Luca et al. [10] suggested that, in addition to the knowledge of the pattern, user’s drawing behavior can be used as a second authentication credential. Currently, they achieved 77% accuracy with 19% FRR and 21% FAR on their 31-participant study. This performance is impressive but barely practical.

One future direction is to improve the performance by developing a more effective recognition technique. In addition, a multi-session training strategy shall be studied since many studies (including EEG [49], voice biometric [58], and face [59]) have shown that it can be used to model within-user variation more effectively. Lastly, a long-term study about consistency of user’s drawing pattern shall be investigated since the performance may automatically improve over time once a user stabilizes his drawing behavior.

Microsoft picture authentication This graphical password approach has been implemented in Windows 8 operating system as an alternative user-device authentication mechanism on multi-touch display devices. In this approach, a user is asked to perform at least three sequential gestures, i.e., tab, line, and circle, on a user choice of background images where a user is required to memorize the set of gestures as well as the positions at which these gestures were performed [11].

As it is a graphical password scheme, memorability is believed to be its advantage over alphanumerical password scheme. However, several studies have indicated that the entropy of user passwords is much smaller than that of theoretical guarantee [11], [12], which due in part to the well-known hotspot problem. In addition, shoulder surfing attack is one shortcoming of this authentication mechanism, similar to other something-you-know authentication schemes. One possible solution to overcome this limitation is to devise a challenge-response based system. For example, instead of using a single set of three sequential gestures on one particular image, the system may prompt a user to set up multiple sequences over multiple images and ask for only a subset of those sequences during verification. The problem related to such an approach would then be how many sequences a user could memorize, or password interference problem, which shall be investigated.

Online signature A finger-drawn signature is another plausible candidate given that signatures are socially and legally accepted authentication credentials. In this scheme, a user will be authenticated by drawing a legitimate signature on touch sensitive display using his fingertip. The system to verify a authenticity of an online signature that is drawn using stylus in controlled environment has long been studied and the verification performance has been reported using the dataset where the signatures are drawn using stylus in controlled environment [17]–[20]. In terms of verification performance when an online signature is drawn on personal mobile devices in an unsupervised setting, with the algorithms proposed in [22], 3% EER was achieved given that a user template is generated from enrolled samples across multiple sessions [21]. The study also showed that the drawing signature has an advantage over the 4-digit PIN in securing users against guessing attacks. In addition, the scheme also has an intrinsic security advantage over share-secret approaches against shoulder surfing attacks as the system also leverages user’s signing behavior in ensuring the identity of a user. One direction for future work of this approach is to quantify the quality of enrolled signatures during the enrollment process, similar to password strength meter, so that the system could inform users with weak signatures and alternately prompt them to enter the new ones.

Multi-touch gesture authentication This approach is an alternative authentication mechanism proposed by Sae-Bae et al. [13]–[15]. The approach is more specific to touch interface as it requires users to directly interact with touch sensitive display using multiple fingers simultaneously, as opposed to interacting with a pointing device. This natural and fluid multi-touch interaction can also be used as a proof of user identity. That is, authentication credential can be drawn from something-
you-know, which is the type of gesture password, as well as something-you-are, which is the movement characteristic of user fingertips that is unique among users. The biometric recognition performance of multi-touch gestures is reported at the rate of 4.37-19.23% EER depending on the type of gestures being performed. In addition, the study has shown that usability of the gestures in terms of ease of use, excitement, pleasantness, and willingness to used is correlated well with verification performance. This is indeed an encouraging result towards usable and secure authentication approach. Another study by Shahzad and Liu has also shown that gestures performed 1-3 fingers, in stead of 5-finger gestures, can also be used for authentication [16].

One direction of future work in this research area is to improve recognition performance by developing a more effective training strategies and recognition method as well as leveraging utilities of UI elements such as background or visual feedback to enhance user consistency and memorability.

III. 2-D SENSING CAMERA INTERFACES

A 2-D camera is another sensor that is widely available on many off-the-shelf device categories including mobile phones, tablets, smart watch, laptop, as well as Google Glass. In addition to the use of sensor as an image/video recorder, it can be used as a communication channel between in devices and users [60]. For example, a user can perform hand gestures to execute various devices’s functionalities [61]–[63]. Regarding authentication methods that make use of 2-D camera as a user interface, a number of approaches have been proposed.

Hand gesture authentication This is one authentication approach that is applicable on 2-D camera. In particular, Fong et al. have proposed to use a hang sign language as a mechanism to input textual password entries on 2-d camera interfaces [23] (See Figure 5 for an example of hand language that can be used to input textual password). In turn, the system can make use of biometric information, i.e., hand shape and behavior characteristics, as an additional authentication factor thereby providing better security against shoulder surfing attack or insiders. Technically, they make use of static images that are captured while a user performs a sign language to verify both a signer’s identity and signed content. The proposed signer verification as well as signed content are evaluated on a four-user set and the best recognition performance of 93.75% is achieved. However, validation of the result on a larger set of users, usability and time it takes to memorize sign language for unfamiliar users has not yet been investigated.

Face recognition This is one of the most well-known authentication mechanism that use 2-D camera as an interaction interface. In addition to airport security application, the system has also been offered as an unlocking mechanism on a number of operating system platforms including Microsoft window, Ubuntu, and android. Security of face recognition as an authentication mechanism Verification performance of face recognition system depends on the context of use. For still face under uncontrolled illumination condition, in Face Recognition Vendor Test(FVRT) 2006 [24], 11% and 13% FRR are measured at 0.1%FAR on two datasets: Notre Dame data set and Sandia. However, in the context of mobile environment, face verification system achieve much lower performance at 10.9% HTER (Half Total Error Rate) [25] due to several factors including facial occlusion, pose and facial expression. Figure 6 presents an example of commercial face locking application products.

However, there is always a tradeoff between false acceptance and false rejection rate especially when the system is employed in uncontrolled usage scenario. Consequently, it could impact security and usability of the system. In addition, the system is vulnerable to several security threats. For example, an attacker may fool the system by presenting a picture or video of the authorized user instead. To counter this attack, liveness detection mechanisms have been proposed. For example, the system may challenge a user to perform some task such as blinking an eye [26] in order to verify liveness of the presented proof (a user’s face) or looking at and following a secret icon [64]. Nevertheless, usability of these preventive mechanisms shall be investigated since user adoption is an important factor for their success.

Fingerprint recognition One plausible biometric authentication approach on this interface is to acquire a fingerprint impression from a camera and use it as a factor to authenticate users [27], [28]. However, one limitation of this approach is that the quality of fingerprint impression taken from a camera is typically lower than the one taken from fingerprint scanners due to several factors. These include variation of lighting conditions, the focus of camera’s lenses, image distortions, and scale and pose variations of a user’s finger while its image is taken. As a result, verification performance can significantly deteriorates. The verification performance of fingerprint recognition with Embedded Cameras on mobile phones has been reported at 4.5% Equal Error Rate (EER) [27] as compared to around 0.02% when fingerprint images are with very high quality [65]. In addition, its usability in terms of verification time, failure
his signature in 3D space while holding the phone. Later, Jing et al. have proposed a similar concept, KinWrite, but using Kinect, or 3-D sensing camera to track a fingertip motion while a user is writing his password [30]. Their method to recognize a user’s 3D signature achieved 100% precision at an average of 77% recall rate on their experiment involving 35 signatures from 18 subjects collecting over a period of five months. Its usability and memorability are however still questionable and shall be investigated.

**Gesture-Based User Authentication** The scheme is proposed by Lai et al. [33]. In this scheme, a user is authenticated by performing a whole-body gesture in front of a 3-D camera. The system then compares and accepts authentication interaction if and only if the input gesture is in close proximity to the enrolled gestures is within a close proximity, i.e., the dissimilarity score between them is lower that the predefined threshold. The cross-session verification performance was reported at 10.5% and 1.5% for the pre-defined S-gesture and user-defined gesture, accordingly [34]. The system was also attested in different scenarios including users wearing coats or carrying bags where the performance degradation is observed. One possible way to improve verification performance is to model intra-user variation across multiple sessions, which have proven successful in other biometric modalities, e.g., ECG [71] and speaker verification [72]. Another possibility is to develop a multi-factor authentication system by incorporating other contextual information, e.g., face and generic body skeleton model, which could be observed simultaneously by the same interface. These are the open challenges to be pursue in future work.

**SignWave** This is the commercial authentication application that has been launched in Airspace store since 2013 [31]. Using this application, users can simply wave their hands in front of the Leap motion controller to gain access to computing devices. The application relies on the set of geometric features derived from a user’s hand to authenticate the user. The application has been well-received by users as it has been downloaded more than a million times. Unfortunately, the security it has offered is currently of the limit and it can be circumvented quite easily [32]. Several factors that could contribute to this issue include, but are not limited to, accuracy and precision of raw information retrieved from the controller, effectiveness of verification algorithm, and sensitivity of the controller itself to lighting and other environmental conditions. These are the topics that shall be studied in order to develop a more robust solution to authenticate users by way of this hand waving interaction. In addition, other visual information like

![Fig. 6: An example of face locking application available on iOS (http://www.wired.com/2012/04/facevault-app-face-recognition/)](image)

### TABLE II: Authentication approaches on touch interfaces and their authentication factors

<table>
<thead>
<tr>
<th>Approaches</th>
<th>Authentication factors</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Textual password on a virtual keyboard</td>
<td>a sequence of alphabet</td>
<td>[6], [7]</td>
</tr>
<tr>
<td>Microsoft picture</td>
<td>a combination of tab, line, and circle gestures and their associative positions on a background image</td>
<td></td>
</tr>
<tr>
<td>Android screen-lock</td>
<td>a dot-connected pattern</td>
<td>[10]</td>
</tr>
<tr>
<td>Online signature</td>
<td>a user’s signature</td>
<td>[17]–[22]</td>
</tr>
<tr>
<td>Multi-touch gesture</td>
<td>a sequence of multi-touch gestures</td>
<td>[13]–[16]</td>
</tr>
</tbody>
</table>

A 3-D sensing camera or depth camera is another emerging interface that allows users to interact with computer in more tangible ways, i.e., more direct reflection of real-world interaction [69]. The 3-D sensing camera or depth camera technology allows depth of objects to be identified, in addition to color information. Using this depth information, object tracking and segmentation algorithm can be more robust and more accurate [70]. Therefore, it has been widely used as a human computer interaction interface. Kinect—a body tracking game controller for Xbox developed by Microsoft, and Leap Motion—a hand tracking computer controller, are the two well-known examples of using a 3-D sensing camera as an interaction interface.

**In-air signature** This is one authentication approach on 3-D sensing camera that has been proposed. In this scheme, users are authenticated by writing their passwords in 3-D space. The system then accepts a user if the writing is closed enough to the enrolled ones, and rejects otherwise. The first variation of this approach make use of an accelerometer embedded on regular smartphone devices [29] as a sensor. A user then sign his signature in 3D space while holding the phone. Later, Jing
hand’s skin-texture and appearance may be used as additional authentication factors to improve system verification accuracy.

V. VOICE INTERFACES

Voice interface is one of human computer interaction platforms that allows users to communicate with the devices using hand free, natural interaction. To enable this type of interaction, the system acquires an audio signal from a user using an embedded microphone and then translate the signal into machine language.

It has been used as one of the main communication channels between users and devices in several wearable device categories including smart-watch; e.g., Samsung Galaxy Gear, Sony SmartWatch, and augmented reality headsets (with head-up display); e.g., Google Glass, Vuzix m100. In addition, it has been embedded in many handset devices, i.e., mobile phone, tablet, laptop, as well as other smart devices, e.g., thermostat [73].

Voice authentication In this approach, a user is authenticated by speaking in normal level. Typically, voice authentication or speaker recognition system are broadly classified into two variations: text-independent and text-dependent speaker verification [35]. The first one is referred to the system in which speaking contexts of a user can be arbitrary [36] where the second one is referred to the system in which a speaking context has to be the same [37]. The system of the first type could prompt a user to speak a randomly assigned context and verify both the context as well as a speaker thereby providing security against replay attack. On the other hand, the system of the second type could use a speaking context as an additional authentication factor thereby providing a better recognition performance resulting in lower both false acceptance and false rejection rates. However, replay attack becomes much more feasible [38].

In terms of verification performance, a wide range of results have been reported due to variation of devices, environmental factors, and verification techniques. For a mobile device dataset collected from 48 enrolled users and 40 imposters, Ram et al. [39] have reported text-dependent verification performance at 7-11% equal error rate, depending on experimental environments, when testing against the samples of the same context spoken by imposters. Note that, each of the samples in this experiment was a spoken ice-cream flavor phrase. For text-independent verification performance, Baloul et al. [40] have reported 1-12% equal error rate, depending on system configuration and number of training sentences, on CMU PDA dataset collected from 16 users. Note that, each of the samples in this experiment was a spoken sentence of 4 to 6 seconds, as opposed to the spoken phrase in the previous study.

In terms of its usability for user authentication applications on personal computing devices, speech or voice authentication is reviewed as an insecure and intrusive choice in some situations (and device categories) [41]. In addition, several other issues could impact security and usability of this authentication approach and they should be taken into consideration prior to deployment of the approach. These include human factors such as Lombard effect, i.e., speakers increase their vocal effort to compensate for signal to noise ratio in noisy condition [39], emotions, vocal organ illness, aging, and level of attention [74]; device and environmental factors such as level of noise background [75], acoustic disturbances like echo, and microphone frequency response; and attack vectors such as computerized imitation [74].

Whisper authentication is another derivative of voice authentication that is less intrusive to users and more resistant to the secret phrase being overheard. With respect to text-independent whisper authentication system, the techniques from regular voice authentication system shall be modified due to intrinsic difference between signal spectra of whispered and normal speech [76]. In particular, Xiaohong and Heming [42] have demonstrated that adaptive fractional Fourier transform cepstral coefficients are a more effective feature set than MFCC coefficients in identifying a speaker from whisper signal. They were able to achieve almost 100% identification accuracy. However, to the best of our knowledge, the study on whisper signal as a biometric trait (or text-independent whisper authentication system) has not been conducted. In addition, a usability study on whisper authentication approaches have never been performed.

VI. EYE TRACKING INTERFACES

Eye tracking interface is one of the latest user interface technology that is expected to be used in consumer electronic devices and smart vehicles in the near future [77], [78]. With this technology, users can hand freely communicate with the devices using gaze interaction. In general, there are two types of hardware for eye trackers: wearable and fix station ones [79]. To accurately capture user’s eye movement, the sensor hardware may include optical sensors, near-infrared emitting diode, and imaging devices.

Gaze authentication With respect to user authentication applications, the idea of using gaze movement as a way user enter a secret to the authentication system has been proposed and experiments to evaluate usability of the system have been conducted. Particularly, in 2007, Luca et al. have proposed and evaluated the gaze entry for PIN authentication using dwell time, look-and-shoot, and gaze gesture interaction techniques [43]. Another variation of the authentication system using gaze interaction is to use gaze entry as a way enter a set of clicked points in a cued-recall graphical authentication [44], [45]. Note that the cued-recall graphical authentication is the authentication system where a user is expected to choose and memorize a certain number of points on background images, one point per image.

Its advantages include security against shoulder surfing attack since this type of interaction is hardly observable by passengers. In addition, as eye movement is one of the behavioral biometric modalities that has been studied [46], gaze authentication systems could utilize this information to provide secondary security layer to users gaze [47]. Moreover, other biometric traits that could be monitored non-intrusively like face or iris could also be used as part of the authentication factors to further strengthen the security and verification performance of the system. However, sensor calibration and sensitivity, user
familiarity with the gaze entry system and usability of the authentication interaction, as well as the problem of separation between intended and unintended gaze interaction are the challenges that shall be addressed [43], [44].

VII. BRAIN COMPUTER INTERFACE

Brain Computer Interface is another communicating channel between users and machine [80] that do not require physical movement of users. Currently, electroencephalographic activity (EEG) is used to measure electrical signals of brain activities. Using this technology, many vendors have set to offer commercial grade BCI products including Neurosky and Intel [81].

Brain wave authentication Motivated by advancement of Brain Computer Interface (BCI) technology, J. Thorpe et al. have envisioned Pass-thought authentication, a two factor authentication method resistant to physical observation [48]. The first factor is thought secret whereas the second one is user’s brain signal that is unique among users.

Later, in 2007, the feasibility of using brain activity for user authentication purpose has been studied by S. Marcel and J. Millán [49]. In particular, they setup an experiment to evaluate biometric performance of such approach by asking 9 users to perform three brain tasks while wearing the electrode cap that can record brain wave signal(EEG). For a given user, the experiment ran for three days; four sessions per day with 5-10 minutes break between each session. They achieved verification performance of 35.5% HTER (Half Total Error Rate) when training and testing samples are drawn from different days and it degrades over time. However, the performance revamp to 12.9% HTER when training samples are drawn from two days, instead of one day. This initial result has shown some potential but needs major improvement both in terms of performance and usability.

Recently, J. Chuang et al. [50] have investigated the scheme using consumer-grade EEG headset devices (see Figure 7), as opposed to clinical grade ones. In this paper, performance, usability, and recall of user authentication using brain wave signal are studied. In particular, they designed 7 different brain tasks and perform a two-session experiment on 15 subjects. Each user is asked to select their secret for each of the four tasks in the first session and is asked to recall it in the second session. In addition, for both session, a user is asked to perform tasks and answer series of usability questions. The result demonstrated feasibility of the scheme using consumer grade hardware by presenting performance improvement over the previous work and on a larger population. However, one challenge for research in this area is to design generic tasks that are both usable and secure. In addition, while an initial investigation of robustness of the scheme against impersonation attack has been conducted [82], the study is limited only to the set of three attackers where the though process of the three attackers might differ from the users due to their native languages and social environments by which they have been surrounded. Future work to address this impersonation attack more intensively is warranted.

VIII. EVALUATION METRICS FOR USER-DEVICE AUTHENTICATION

Since the ultimate goal of an authentication mechanism is to prevent user’s information from an unauthorized access, security benefit against various threat models is one of the most important criteria in evaluating authentication mechanisms. However, similar to any other security mechanism, an authentication mechanism will only be useful if it is adopted by users. Therefore its usability aspects need to be investigated simultaneously along with its security potential. With respect to evaluation framework for web based authentication mechanism, a comprehensive list to assess security, usability, and deployability is proposed by Bonneau et al. [5]. However, in contrary to their paper, this paper focuses on user-device authentication. That is, the authentication process is expected to take place locally and with hardware sensor readily available on the device. Therefore, we do not consider deployability properties and other security properties that are more related to remote or web authentication mechanisms as oppose to the local ones.

The summary of security and usability criteria that shall be used to evaluate different types of authentication mechanisms depending on which factors are used to authenticate users, derived from their list is as follow.

A. Security

• Physical observation: An attacker in this threat model could observe including record video of authorized authentication interaction while it is being performed before attempting to be authenticated by the system. Security against threat models is becoming more important features of authentication mechanisms due to the change in usage pattern of computing devices. First, many devices are often used in a social context and in public spaces. In addition, users are surrounded by not only public surveillance systems but also handheld recording devices. In such situations, the assumption that users can enter a password or other shared-secret credentials to the devices privately and securely is no longer hold. One direction for an knowledge-based authentication mechanism to prevent this attack is to use...
challenge-response approaches where a user only presents partial information about the secret at the authentication time. However, the benefit typically come at the cost of reduced usability approaches. Another direction is to also rely on other factors to authenticate users.

- **Targeted impersonation**: An attacker in this threat model attempts to impersonate a victim by using knowledge about personal information. For example, it is known that many people use one form of their date of birth as their PINs. Therefore, PIN based authentication is highly vulnerable to this attack. In addition, some graphical approaches that allow users to select background images of their choices and perform some interaction on them is also vulnerable to this attack since many will choose their family photos and perform interaction on faces of their family members.

- **Random guessing** (throttled or unthrottled): An attacker in this threat model is the one who does not have an access to information about authentication credential and attempts to perform authentication interaction in order to gain a legitimate access to the system. This is one of the most common threat since more sensitive information are being stored on or being access from personal computing devices. This information is another motivation for devices in their theft which could lead to several malicious activities, e.g., financial fraud, identity theft, data loss. The

- **Cross-site attack**: An attacker in this threat model attempts to use a compromised credentials from one system on another system in order to get an authorized access. This is a particularly serious threat for the authentication system that stored physiological biometric trait without appropriate template protection mechanism, since users can not easily replace them once they are compromised. In addition, any knowledge-based authentication factor that are widely used across different sites is also vulnerable to this attack since it is known that users often use them repeatedly across the sites. One countermeasure to handle this attack is for the developers to deploy an appropriate hashing technique to protect privacy of the stored credential.

**B. Usability**

- **Memorability**: This characteristic refers to the amount of cognitive burden imposes to users in order to use the system including knowledge interference that could possibly happen when users have to use multiple credentials for different systems or devices. This is an important usability matrix to be considered particularly for the mechanisms that rely on share secret between users and systems as one of the authentication factors.

- **False rejection**: This characteristic refers to the capability of system to verify and grant access to genuine users when presenting legitimate authentication credentials.

Similar to memorability for shared secret authentication approaches, this is a particularly important property for the approach that relies on biometric information as one of the authentication factors. That is, in a biometric authentication, a user who presents a biometric trait that its representative (or feature set) is within a close proximity to that of the enrolled template will be granted an access. As a result, there are two types of error that could occur: false acceptance and false rejection. While the first one could cause a security concern, the latter one could lead to either an increase in the number of authentication attempts or user rejection and temporary lockout which both cause usability issues. Therefore, for an ideal system, these two rates should be kept very low. For many biometric modalities, there are several environmental factors that could affect the verification performance of the system. These include but are not limited to level of noise background (for voice), lighting condition (for image or video), humidity (for capacitive sensor like fingerprint).

- **Ease of use**: This characteristic refers to user effort and efficiency in using an authentication mechanism including ease of learning it. Since there is an implicit economic cost involved when users decide on security issues [83], this is one of the most important property for an authentication system to be adopted.

Typically, authentication factors derived from an authentication interaction can be either shared knowledge, users' biometric traits, or the combination of both, where each of the factor has its own set of advantages and disadvantages. Based on the type of authentication credential, the evaluation of authentication approaches that are reviewed in this paper is summarized and presented in Table III. Nothing that the security and usability evaluation criteria are the ones mentioned above. Regarding the fusion approaches, the more credentials being used by the system, the harder the system to be circumvented. However, this will also increase the chance of genuine and honest users being rejected. Therefore, it is always important to strike a balance between security and usability trade-off.

**IX. DISCUSSION AND FUTURE WORK**

The capability of these emerging interfaces in detecting a large amount of user interaction information has created a unique opportunity to redesign and develop authentication mechanisms that could provide intrinsic security to the user. Specifically, much research have shown that it is feasible to design the user authentication mechanisms that use this information to verify the identity of a user based on who-they-are, in addition to what-they-know. This is one interesting direction to enhance the security of authentication mechanisms.

In addition, many off-the-shelf computing devices that are entering the consumer market have equipped with multiple interfaces to enable users to communicate with the device using an array of interaction. Therefore, one research direction is to form a multi-factor authentication system by simultaneously using multiple interfaces to capture different authentication factors while users are performing an authentication interaction.
For example, with gaze or eye movement as an authentication interaction, geometry and appearance of a user’s face could easily be capture all together using embedded 2-D and 3-D cameras and use as additional authentication factors. As such, it would be much harder to circumvent the system.

Moreover, studies have shown that each of the authentication approaches has its own limitations in terms of environmental conditions it can operate on and has its own weaknesses on different security threat models. And for wearable and augmented devices, their use case scenarios are more diverse and by which environment and situation the users will be surrounded are not restricted. Therefore, another interesting direction for future work is to design a mechanism that allows users to enter the same secret in multiple ways depending on availability of hardware sensors on the devices, surrounding environments, and threat models at the moment. For example, in the Google Glass, users have multiple ways to communicate with the device including voice command, hand gestures, and tab sliding. Their suitability depends on the situation. For example, voice channel is not an appropriate way to be used in the places that people are supposed to be quiet and it would be less effective in a noisy place. In addition, hand gestures are not robust to perform in a dark place or clutter background. Therefore, having a mechanism that allows user to operate in multiple ways would increase overall usability of the system.

Lastly, user effort and accessibility when authenticating to the system are other important factors to be evaluated, in addition to the system provided to the user, since any security tool would be useful only if it is adopted by the user. Also, memorability is another issue to consider if the something-the user-know is also part of the authentication credential. Therefore, these factors shall be considered when designing and evaluating authentication alternative mechanisms.

### REFERENCES


[77] “General Motors to fit eye-tracking technology that reveals when a driver is not paying attention to the road.” http://www.dailymail.co.uk/sciencetech/article-2740588/, 2014. [Online; accessed 09-Sep-2014].


