# A Survey of Human-Computer Interaction (HCI) & Natural Habits-based Behavioral Biometric Modalities for User Recognition Schemes

Sandeep Gupta[a,c], Carsten Maple[b], Bruno Crispo[c], Kiran Raja[d], Artsiom Yautsiukhin[a], Fabio Martinelli[a]

[a]Istituto di Informatica e Telematica (IIT), Consiglio Nazionale delle Ricerche (CNR), Pisa, Italy
[b]University of Warwick, Coventry, UK
[c]Department of Information Engineering & Computer Science (DISI), University of Trento, Italy
[d]Norwegian University of Science and Technology (NTNU), Norway

**Abstract**

The proliferation of Internet of Things (IoT) systems is having a profound impact across all aspects of life. Recognising and identifying particular users is central to delivering the personalised experience that citizens want to experience, and that organisations wish to deliver. This article presents a survey of human-computer interaction-based (HCI-based) and natural habits-based behavioral biometrics that can be acquired unobtrusively through smart devices or IoT sensors for user recognition purposes. Robust and usable user recognition is also a security requirement for emerging IoT ecosystems to protect them from adversaries. Typically, it can be specified as a fundamental building block for most types of *human-to-things* accountability principles and access-control methods. However, end-users are facing numerous security and usability challenges in using currently available knowledge- and token-based recognition (*i.e., authentication and identification*) schemes. To address the limitations of conventional recognition schemes, *biometrics*, naturally come as a first choice to supporting sophisticated user recognition solutions. We perform a comprehensive review of touch-stroke, swipe, touch signature, hand-movements, voice, gait and footstep behavioral biometrics modalities. This survey analyzes the recent state-of-the-art research of these behavioral biometrics with a goal to identify their attributes and features for generating unique identification signatures. Finally, we present security, privacy, and usability evaluations that can strengthen the designing of robust and usable user recognition schemes for IoT applications.

*Keywords:* Internet of Things (IoT), User Recognition, Behavioral Biometrics

## 1. Introduction

IoT ecosystems, integrating smart sensors, actuators, advanced communications, efficient computation, and artificial intelligence, have the power to transform the way we live and work. Almost every business vertical has started to embrace IoT technology [1]. This includes sectors as diverse as automotive, energy, entertainment, education, food, finance, healthcare, and transportation where smart, integrated systems are delivering improved quality of life and resource efficiency by providing security-sensitive services via IoT applications. Bera et al. [2] reported that user authentication, access control, key management, and intrusion detection are essential requirements to prevent real-time data access directly from the IoT-enabled smart devices that are deployed in IoT ecosystems. Studies have indicated that application-layer attacks in the IoT are particularly complex to detect and deflect [3, 4]. Ultimately, any security breach of IoT ecosystems has the potential for profound consequences on consumers and society [5]. Therefore, robust and usable *Authentication*, *Authorization* and *Accounting* (`AAA`) mechanisms for applications bridging humans and IoT ecosystems, which can be specified as IoT Applications, are critical for maintaining *confidentiality*, *integrity*, *availability* (`CIA`) in the system.

Many IoT ecosystems still rely on traditional Personal Identification Numbers (PINs), passwords, and tokens based user recognition mechanisms [6]. This is despite, users facing both security and usability challenges in using these conventional *(knowledge- and token-based)* recognition schemes [7, 8]. Further, the decision process in conventional authentication mechanisms is usually binary [9]. PINs and passwords can be easily guessed, shared, cloned, or stolen [10]. Conventional authentication schemes are also prone to a wide range of common attacks [11], such as dictionary-, observation- and replay-attacks. Weak passwords remain the major cause of botnet-based attacks, such as Mirai, on huge numbers of IoT systems [12]. Additionally, they possess several usability issues [13], such as placing overwhelming cognitive load on users and ergonomic inefficiencies for newer IoT end-points. As such, human-to-things recognition schemes for IoT ecosystems require rethinking, with behavioral biometrics providing an appropriate alternative to overcoming the drawbacks present in conventional authentication schemes.

This article presents a comprehensive review of *touch-stroke, swipe, touch signature, hand-movements, voice, gait* and *footstep* behavioral biometric modalities for designing user recognition schemes in emerging IoT ecosystems. The motivation for this particular selection of modalities is provided by the current focus of academic research, and the industrial trend towards human-computer interaction (HCI) and

---

*Email address:* `sandeep.gupta@ex-staff.unitn.it` (Sandeep Gupta)

natural habits-based behavioral biometrics-based recognition schemes. For instance, *ViewSonic* and *Namirial* partnered to deliver a behavioral biometric eSignature solution that includes the behavioral biometric of handwritten signatures to boost electronic signature security and reliability [14]. *Banking sectors* are investigating characteristics including touch-stroke dynamics to generate a trusted user profiles for distinguishing between normal and unusual user behavior, as a means to detect fraudulent users [15]. Other companies, such as *BehavioSec* [16] and *BioCatch* [17] are leveraging behavioral biometrics, including swipe or touch gestures, typing rhythm, or the particular way an individual holds their device, to offer enterprise-scale security solutions for continual and risk-based authentication or fraud detection, for example. Electronic payment card providers are investigating behavioral biometrics for cutting-edge payment systems of the future [18]. A study of biometrics to achieve intelligent, convenient, and secure solutions for smart cities and smart transportation are presented in [19] and [20], respectively. Sensor-based activity recognition [21], such as gait, can be used to verify commuters through their walking patterns, thereby replacing the need for a travel pass to access public transportation. *NEC Corporation* and *SITA* have collaborated to roll out a walk-through, contactless digital identity solution for airports leveraging their biometric identity management platform to facilitate a non-intrusive method of identity verification [22]. So large is the potential that the market study forecasts that by 2025 behavioral biometrics market will reach 3.92 Billion [23].

### 1.1. Objectives and survey strategy

The objective of this article is to survey HCI and natural habits-based biometrics that can be utilized by researchers and engineers to design uni-modal or multi-modal user recognition schemes (leveraging concepts such as implicit, continuous, or risk-based [9]) for security-sensitive applications, thus, safeguarding IoT ecosystems.

Table 1 lists previous surveys related to the behavioral biometric modalities covered in this article.

Table 1: Earlier behavioral biometrics surveys

| **Ref** | **Year** | **Contributions** |
|---|---|---|
| Yampolskiy and Govindaraju [24] | 2008 | This survey presented a classification of behavioral biometrics based on skills, style, preference, knowledge, motor skills, or strategy applied by humans. |
| Meng et al. [25] | 2015 | This survey covered the development of biometric user authentication techniques on mobile phones. And, presented a study of voice, signature, gait, behavior profiling, keystroke and touch dynamics behavioral biometrics. |
| Alzubaidi and Kalita [26] | 2016 | This survey investigated authentication of smartphone users based on handwaving, gait, touchscreen, keystroke, voice, signature and general profiling behavioral biometrics. |
| Oak [27] | 2018 | This survey analyzed persons' behavior, such as keystroke dynamics, mouse dynamics, haptics, gait, and log files, for their designing persistent security solutions. |
| Dang et al. [28] | 2020 | This survey focused on Human activity recognition (HAR) for designing context-aware applications for emerging domains like IoT and healthcare by analyzing sensor- and vision-based behavioral patterns. |

| Ref. | Year | Contributions |
|---|---|---|
| Stylios et al. [29] | 2020 | This survey presented the classification of behavioral biometrics technologies. It reviewed behavioral traits like gait, touch gestures, keystroke dynamics, hand-waving, behavioral profile, power consumption, for continuous authentication for mobile devices. |

In this survey, we first elucidate attributes and features of behavioral biometric modalities that can be acquired from smart devices equipped with motion sensors, touch screens, and microphones or by external IoT sensors or nodes in an unobtrusive manner. We discuss the methodologies, classifiers, datasets, and performance results of recent user recognition schemes that employ these behavioral biometrics modalities. We then present *security*, *privacy*, and *usability* attributes with regard to the `CIA` properties in human-to-things recognition schemes. Ultimately, the challenges, limitations, prospects, and opportunities associated with behavioral biometric-based user recognition schemes are presented.

### 1.2. Article Structure

The article is structured as follows: *Section 2* discusses behavioral biometrics, sensors, human-to-things recognition mechanisms and performance metrics. *Section 3* elicits attributes and features of *touch-stroke, swipe, touch signature, hand-movements, voice, gait,* and *footstep* modalities that can be exploited for designing user recognition schemes. *Section 4* presents the state-of-the-arts of user recognition schemes based on modalities discussed in Section 3. *Section 5* presents a discussion on security, privacy, and usability of behavioral biometric-based user recognition schemes. *Section 6* discusses the open challenges and limitations that deserve attention together with prospects and opportunities for evolving and designing behavioral biometric-based human-to-things recognition schemes. *Section 7* concludes the article.

## 2. Background

Despite many advancements in recent years, human-to-things recognition (identification and authentication) remains a challenge for emerging IoT ecosystems [30]. Evidently, with improvements in sensors technology, the opportunity to evolve behavioral biometric-based human-to-things recognition schemes has increased significantly.

### 2.1. Behavioral biometrics

Behavioral biometrics involve human behavioral characteristics or activity patterns that are measurable and uniquely identifiable and so can be designed into user recognition schemes. Typically, behavioral biometric modalities can be considered according to persons' skills, style, preference, knowledge, motor-skills, or strategy while they interact with an IoT application [24]. The categories that can be derived are 1) authorship; 2) HCI; 3) indirect HCI; 4) motor skills; and 5) natural habit, based on various information extracted or gathered from a person. These categories are summarised in Figure 1.

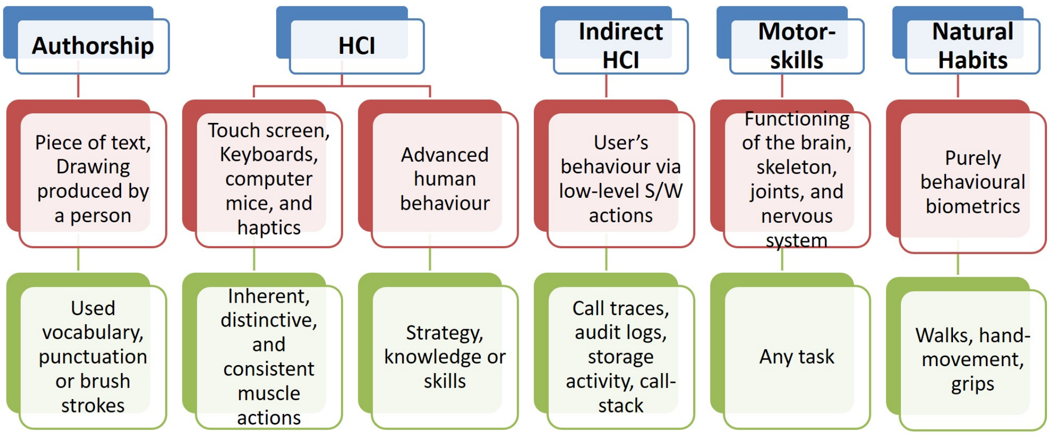


Figure 1: A categorization of behavioral biometrics [24]

- *Authorship-based biometrics* involves verifying a person by observing peculiarities in their behavior. This includes the vocabulary used, style of writing, punctuation, or brush strokes, occuring in their writings or drawing [31].

- *HCI-based biometrics*, exploits a person's inherent, distinctive, and consistent muscle actions while they use regular input devices, such as touch-devices, keyboards, computer mice, and haptics [32]. Furthermore, it leverages advanced human behavior involving knowledge, strategies, or skills exhibited by a person during interaction with smart devices.

- *Indirect HCI-based biometrics* may be considered as an extension of the second category. It considers a person's indirect interaction behavior, by monitoring low-level computer events (e.g., battery usage) [33], stack traces [34], application audit [35], or network traffic logs [36], or mutual interaction analysis (e.g., completely automated public Turing test to tell computers and humans apart - *CAPTCHA*) [37].

- *Motor-skills based behavioral biometrics* can be described as the ability of a person to perform a particular action using muscle movements [38]. These muscle movements are produced as a result of coordination between the brain, skeleton, joints, and nervous system that differs from person to person [39].

- *Natural habits-based biometrics* constitute purely behavioral biometrics measuring persistent human behavior such as gait [40], hand-movement [41], swipe [42], grip [43], and footstep [44].

### 2.2. Sensors

The rapid evolution of system-on-chip (SoC) and wireless technologies play a vital role in evolving smarter, smaller, accurate, and efficient sensors for behavioral biometric data acquisition. Table 2 describes sensors that can be integrated into smart devices and portable IoT devices for acquiring behavioral biometric modalities covered in Section 3.

Table 2: Sensors for acquiring behavioral biometric modalities

| **Category** | **Sensor description** | **Sensor Type** |
|---|---|---|
| Position | Position sensors can be linear, angular, or multi-axis. It measures the position of an object that can be either relative in terms of displacements or absolute positions. | Proximity sensor, Potentiometer, Inclinometer |
| Motion, Occupancy | Motion and occupancy sensors detect movement and presence of people and objects, respectively. | Electric eye, RADAR, Depth Camera |
| Velocity, Acceleration, Direction | Velocity sensors can be linear or angular. It measures the rate of change linear or angular displacement. Acceleration sensors measure the rate of change of velocity. Magnetometer estimates the device orientation relative to earth's magnetic north. Gravity sensor indicates the direction and magnitude of gravity. | Accelerometer, Gyroscope, Magnetometer, Gravity sensor |
| Pressure | Pressure sensors detect force per unit area | Barometer, bourdon gauge, piezometer |
| Force | Force sensors detect resistance changes when a force, pressure, or mechanical stress is applied. | Force gauge, Viscometer, Tactile sensor (Touch sensor), Capacitive touchscreen |
| Acoustic, Voice | Acoustic sensors measure sound levels transform it into digital or analog data signals. | Microphone, geophone, hydrophone |

IoT endpoints (devices) can provide position, orientation, or other motion-based measurements to determine unique and finite *hand micro-movements*. These 3-D space measurements can describe device positioning and movement while users interact. Similarly, acoustic, pressure, motion, or occupancy sensors can be used for acquiring behavioral biometric modalities such as *voice*, *gait*, or *footstep* for user recognition. Touch screens can be utilized to acquire *touch-stroke*, *swipe*, or *touch-signature* data.

### 2.3. Human-to-things recognition process

ISO2382-2017 [45] specified biometric recognition or biometrics as an automated recognition of individuals based on their biological and behavioral characteristics. ISO2382-2017 mentioned that the use of 'authentication' as a synonym for "biometric verification or biometric identification" is deprecated; the term biometric recognition is preferred. Thus, human-to-things recognition can be a generic term encompassing automated *identification* and *verification* of individuals in the context of IoT applications.

- According to ISO2382-2017 [45], an identification process is a *one-to-many comparison* decision to determine whether a particular biometric data subject is in a biometric reference database. Identification systems can be employed for both negative recognition (such as preventing a single person from using multiple identities) or positive recognition for authentication purposes.
- Similarly, ISO2382-2017 [45] defines a verification process as a comparison decision to determine the validity of a biometric claim in a verification transaction. Thus, a verification process is a *one-to-one comparison* in which the biometric probe(s) of a subject is compared with the biometric reference(s) of the subject to produce a comparison score. Generally, a verification system requires a labeled claimant identity as an input to be compared with the stored templates (e.g., biometrics templates) corresponding to the given label, to assert the individual's claim. Often, verification systems are deployed for positive identification to prevent systems from zero-effort impostors and illegitimate persons.

### 2.4. Performance metrics

In a biometric system designed to distinguish between a legitimate user or an impostor, there can be four possible scenarios. These are derived from the person being legitimate or not, and being (correctly or incorrectly) identified as legitimate or not. These are termed true acceptance ($TA$) or false rejection ($FR$) and true rejection ($TR$) or falsely acceptance ($FA$) [46]. We describe the most commonly used indicators for the performance evaluation of biometric systems.

- **True Acceptance Rate (TAR):** This is the ratio of $TA$ legitimate user attempts to the overall number of attempts ($TA + FR$). A higher TAR indicates that the system performs better in recognizing a legitimate user.
- **False Rejection Rate (FRR):** This is the ratio of $FR$ legitimate user attempts to the overall attempts ($TA + FR$). FRR is a complement of TAR and it can be calculated as `FRR = 1 - TAR`. ISO/IEC 19795-1:2006 [47] also denote the term FRR as False Non-Match Rate (FNMR).
- **False Acceptance Rate (FAR):** This is the ratio of $FA$ impostor attempts to overall attempts ($FA + TR$). A lower FAR means the system is robust to impostor attempts. ISO/IEC 19795-1:2006 [47] also specified the term FAR as False Match Rate (FMR).

- **True Rejection Rate (TRR):** This is the ratio of $TR$ attempts of impostors to all overall attempts ($FA + TR$). TRR is the complement of FAR and can be calculated as `TRR = 1 - FAR`.
- **Equal error rate (EER):** It is the value where both errors rates, `FAR` and `FRR`, are equal (i.e., `FAR = FRR`).
- **Accuracy:** The ratio of ($TA + TR$) to ($TA + FR + TR + FA$).
- **Receiver- or Relative-Operating Characteristic (ROC):** ROC plot is a visual characterization of trade-off between `FAR` and `TAR` [47]. In simple terms, this is a plot between correctly raised alarms against incorrectly raised alarm. The curve is generated by plotting the `FAR` versus the `TAR` for varying thresholds to assess the classifier's performance.
- **Detection Error Trade-off (DET) Curve:** A DET curve is plotted using `FRR` and `FAR` for varying decision thresholds. To determine the region of error rates, both axes are scaled non-linearly [47]. Deviation- or logarithmic scales are the most commonly used scales in such graphs.

## 3. Behavioral Biometric Modalities' Attributes and Features

This section presents the attributes and features of behavioral biometric modalities that can be exploited for conceptualizing and designing human-to-things recognition schemes. In particular we examine behavioral biometric modalities based on HCI and natural habits that can be collected with no explicit user input using users' smart devices, e.g., smart devices, smartwatches, etc., or external IoT sensors/nodes, e.g., pressure sensors, camera, etc.

### 3.1. Touch-strokes dynamics

Touch-strokes can be described as touch sequences registered by a touchscreen sensor while users navigate on touchscreen-based smart devices using their fingers [48]. Studies have shown that human musculoskeletal structure can produce finger movements that can differ from person to person [49]. Thus, a unique digital signature can be obtained from individuals' touch-points or keystrokes collected using built-in touch sensors available in smart devices. Commonly, touch-stroke features can be categorized as spatial, timing, and motion features [50].

#### *3.1.1. Spatial features*

Spatial features for touch-stroke involves physical interactions between a user fingertip and a device touchscreen surface that can be acquired when a touch event is triggered. Subsequently, a cumulative distance, i.e., a sum of lengths computed from all the consecutive touchpoints in the 2-D space, and speed, i.e., cumulative distance divided by total touch-time, can be derived from touch events [51]. Commonly used spatial features are touch positions, time-stamp, touch size, and pressure [52, 53].

### 3.1.2. Timing features

The touch-stroke timing features generation method can utilize dwell (*press or hold*) and flight (*latency*) time. *Dwell time* can be defined as the time duration of a touch-event of the same key and *flight time* can be defined as the time interval between the touch events of two successive keys. These features are directly proportional to the number of touches on the touch-screen. As an example, Figure 2 illustrates 30 features containing 8-*Type0* dwell time features and 22-*Type1* to *Type4* flight time features that can be extracted from the 8 touch-sequence [54].

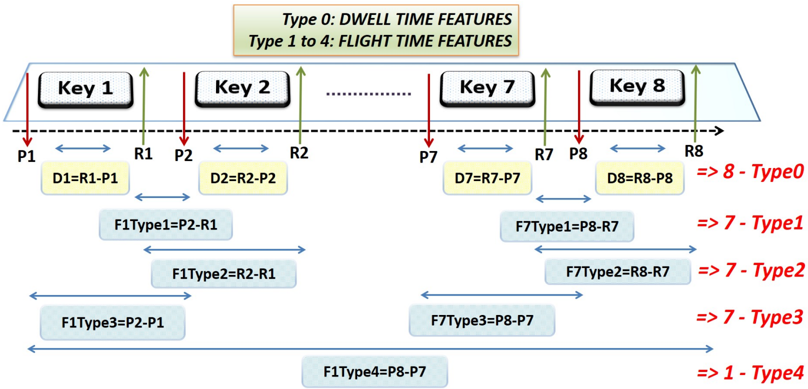


Figure 2: Commonly used duration based touch-strokes timing features

The touch-stroke timing features generation method can also utilize different key-touch duration as illustrated in Figure 3. The shortest feature-length can be termed as uni-graph, which is the timing feature extracted by taking the touch event timestamp values of the same key [55]. The timing features extracted from two, three, or more keys are termed as di-graph, tri-graph, and n-graph, respectively.

### 3.1.3. Motion features

Motion features can be acquired using motion sensors, such as Accelerometer, Gyroscope, Magnetometer, or gravity sensors that are available in most smart devices. Each touch event normally inflicts some movements or rotations that can be registered to generate a unique user authentication signature [56]. However, these motion features can be associated better for other user behaviors like hold- and pick-up movement [57].

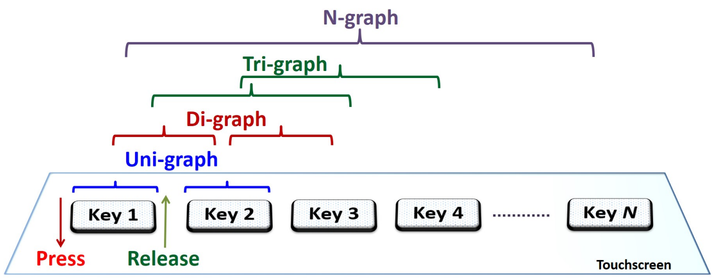


Figure 3: Graph based touch-strokes timing features

### 3.2. Swipe

Swipe can be defined as a finite touch-events sequence that occurred as a result of users touching a smart device's touchscreen with their finger. Smart devices provide APIs to get touch coordinates, velocity, and pressure data for each touch-point [58].

Some of the spatial features that can be extracted from a swipe action are the touch-points timestamp, x- and y-coordinates, velocity, and acceleration. Acceleration for each touch-point can be computed mathematically, from velocity data. The touch pressure of each touch-point determines how hard the finger was pressed on the screen, and what was the touch size. Also, trajectory length, duration, average velocity, average touch-size, start and end touch coordinates can be derived from a swipe data [59, 60]. Additionally, statistical features, such as min, max, average, standard deviation, variance, kurtosis, and skewness can be computed from each 2-D touch sequence, i.e., position, velocity, acceleration, and pressure, acquired for a swipe action [61].

### 3.3. Touch Signature

Touch signature, i.e., a person signing on smart devices' touchscreen using their finger or stylus, is similar to a handwritten signature. Although, a touch signature can utilize the features that are extracted for a swipe gesture to generate a unique identification for users specified in Section 3.2.

Typically, touch signature features can be classified as global and local features [62]. Global features include total writing time, number of strokes, and signature size. Local features include local velocity, stroke angles, etc., computed at an instance of time or for a short duration. Some of the statistical features that can be extracted for touch signature are minimum, maximum, and mean of speed, acceleration, pressure, and size of the continuous strokes [63]. Further, for each stroke in a touch signature, touch-duration, segment direction, log curvature radius, stroke length to width ratio can be extracted [64, 65].

Touch-duration can be utilized for finding similarity between touch signatures of a person. The difference between the two touch-duration sequences ($T_{difference}$) can be computed using Equation 1. $T_s(n)$ and $T_r(n)$ are touch-duration of $n^{th}$ touch sequence, respectively that are obtained from two touch signatures of a person.

$$T_{difference} = \sum_{n=1}^{N} |T_s(n) - T_r(n)| \tag{1}$$

The direction ($\theta_i$) of i-th segment having coordinates ($x_i, y_i$; $x_{i+1}, y_{i+1}$) can be calculated using Equation 2.

$$\theta_i = arctan\left(\frac{y_{i+1} - y_i}{x_{i+1} - x_i}\right) \forall\ i = 1\ to\ N \tag{2}$$

After decomposing the signature into multiple strokes, Lognormal velocity distribution $v_i(t)$ of $i^{th}$ stroke for a given starting time ($t_{0i}$), stroke-length ($D_i$), logtime delay ($\mu_i$) and logresponse time ($\sigma_i$) can be obtained using Equation 3.

$$|v_i(t)| = \frac{D_i}{\sqrt{2\pi}\sigma_i(t - t_{0i})} exp(-\frac{(ln(t - t_{0i}) - \mu_i)^2}{2\sigma_i^2}) \tag{3}$$

### 3.4. Hand Movements

Hand movements can be defined as a finite trajectory in 3-D space for gestures like hold, upward, downward, or snap while users perform a particular activity using their smart devices. For a user's hand-movement action, unique user-identification-signature can be generated from collected $X$, $Y$, $Z$, and $M$ coordinates. In this process, $X$, $Y$, and $Z$ streams can be collected using sensors such as Accelerometer, Gyroscope, Magnetometer, or Gravity sensors, available in smart devices. Whereas, magnitude stream can be derived mathematically, from each sample ($X$, $Y$, $Z$) using Equation 4.

$$M = \sqrt{(X^2 + Y^2 + Z^2)} \tag{4}$$

Where, $M$ is the magnitude and $X$, $Y$, and $Z$ are the X, Y, and Z coordinates obtained from each sensor sample.

Univariate statistical features can then be extracted from each raw stream that aid to reduce the dimensionality of raw data and improve the signal-to-noise ratio [41]. Some of the statistical features, such as *min* (minimum value), *max* (maximum value), *mean* (average value), *standard deviation* (variation from the mean value), *skewness* (measure of the distortion or asymmetry), *kurtosis* (measure of the tailedness), etc., for a dataset ($S$) containing $N$ values can be computed using Equations 5.

$$
\begin{aligned}
Minimum\ (Min) &= \min_{i=1}^{N} S_i & Standard\ Deviation\ (\sigma) &= \sqrt{\frac{\sum_{i=1}^{N}(S_i - \mu)}{N}} \\
Maximum\ (Max) &= \max_{i=1}^{N} S_i & Kurtosis\ (k) &= \frac{\frac{1}{N}\sum_{i=1}^{N}(S_i - \mu)^4}{\sigma^4} \\
Mean\ (\mu) &= \frac{1}{N}\sum_{i=1}^{N} S_i & Skewness\ (s) &= \frac{\frac{1}{N}\sum_{i=1}^{N}(S_i - \mu)^3}{\sigma^3}
\end{aligned}
\tag{5}
$$

### 3.5. Voice

Speech processing can be a challenging task as people have different accents, pronunciations, styles, word rates, speed of speech, speech emphasis, accent, and emotional states. Typically, a voice-based authentication system can be either text-dependent or text-independent. Figure 4 illustrates speech processing methods encompassing speaker identification, speaker detection, and speaker verification [66].

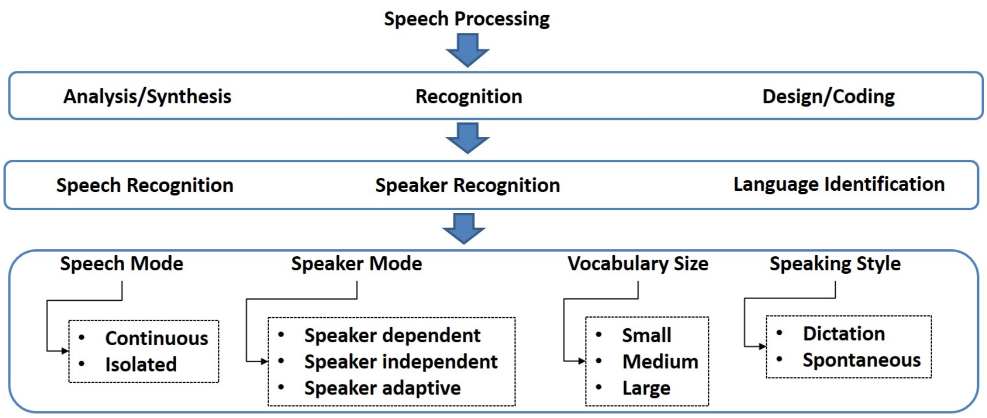


Figure 4: An overview of speech processing [66]

Voice biometrics exploit human speech parametrization or pattern matching/scoring methods to generate a unique identification signature. Human speech generation involves the lungs, vocal cords, and vocal tracts [67]. When a person speaks, the air expels from the lungs passing through the vocal cords that dilate or expand allowing the airflow to produce unvoiced or voiced sound. Subsequently, the air is resonated and reshaped by the vocal tract that consists of multiple organs such as the throat, mouth, nose, tongue, teeth, and lips. The vocal cord's modulation, interaction, and movement of these organs can alter sound waves and produce unique sounds for each person. For a sound, the phoneme is known as the smallest distinctive unit sound of a speech [68] and pitch can be referred to as a fundamental frequency [69]. Each phoneme sound can be explained as airwaves produced by the lungs that are modulated by the vocal cords and vocal tract system.

Speech parametrization transforms a speech signal into a set of feature vectors, such as Mel Frequency Cepstral Coefficients (MFCCs), mean Hilbert envelope coefficients (MHEC) [70], Power Normalized Cepstral Coefficients (PNCCs) [71], and non-negative matrix factorisation (NMF) [72]. MFCCs are widely used parametric features for automatic speech and speaker recognition systems [73]. A Mel is a unit of pitch [74]. The sound pairs that are perceptually equidistant in pitch are separated by an equal number of Mels. The mapping between frequency in Hertz and the Mel scale is linear below 1000 Hz and logarithmic above 1000 Hz. The Mel frequency $m$ can be computed from the raw acoustic frequency.

$$mel(f) = 1127ln(1 + \frac{f}{700}) \tag{6}$$

To extract MFCCs, first the voice signal is pre-emphasized using a first-order high-pass filter to boost the high frequencies energy. The next step involves windowing that can be performed using the Hamming function to extract spectral features from a small window of speech. Afterward, Fast Fourier Transform (FFT) is applied to extract spectral information from the windowed signal to determine the amount of energy at each frequency band. For computing MFCCs, filter banks are created with 10 filters spaced linearly below 1000 Hz, and the remaining filters spread logarithmically, above 1000 Hz collecting energy from each frequency band. After taking the *log* of each of the mel spectrum values. Finally, Inverse Fast Fourier Transform (IFFT) is applied extracting the energy and 12 cepstral coefficients for each frame.

Pattern matching/scoring methods involves probabilistic modeling (e.g., Gaussian Mixture Model (GMM) [75], Hidden Markov Models (HMMs) [76], Joint factor analysis (JFA), i-vectors [75]), template matching (e.g., vector quantization, nearest neighbor) and deep neural network trained on various combinations of i-vectors, x-vector, feature-space maximum likelihood linear regression (fMLLR) transformation [75] or Gabor filter (GF) [77]. I-vectors are low-dimensional fixed-length speaker-and-channel dependent space that is a result of joint factor analysis [78]. For extremely short utterances, i-vectors based approaches can provide an effective speaker identification solution using different scoring methods like cosine distance or probabilistic linear discriminant analysis (PLDA). In an x-vector system, DNN is trained to extract the speaker's voice features, and the extracted speaker embedding is called x-vector [79].

### 3.6. Gait

Human gait is the defined as the manner and style of walking [80]. Gait can be characterized by its cadence that is measured as the number of steps per time unit. Typically, a person's gait varies during different activities, e.g., walking, running, hopping, ascending, or descending, etc. [81]. A gait cycle, illustrated in Figure 5,

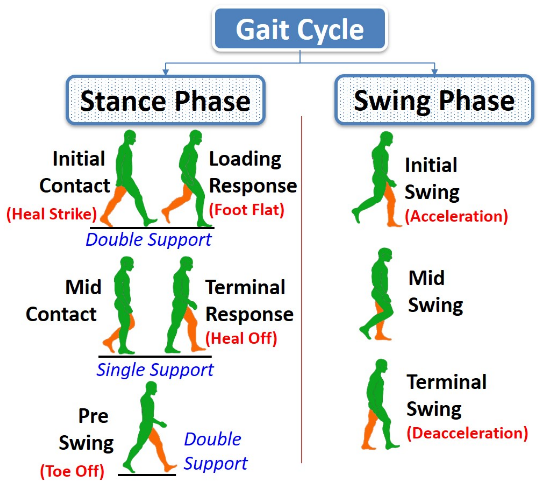


Figure 5: An illustration of a gait cycle

consists of two primary phases: stance and swing [82]. The stance phase is the time-period during which feet are on the ground, constitutes approximately 60% of the gait cycle. The swing phase is the time-period during which the foot is in the air, constitutes the remaining 40% of the gait cycle. A stance phase can be further divided into 1) initial-contact and loading-response, 2) mid-contact and terminal-response, and, 3) Pre-swing. Similarly, a swing phase can be divided into 1) initial, 2) mid, and 3) terminal swing [83]. Using these parameters, both time-based and spatial features can be extracted as indicated in Table 3.

Table 3: Gait features

| # | **Spatial** | **Time** |
|---|---|---|
| 1. | Stride length (cm) | Duration of step (milli sec) |
| 2. | Step length (cm) | Stride duration (milli sec) |
| 3. | Stride width or base of support (cm) | Stance phase (milli sec) |
| 4. | Internal/External Angle (deg) | Swing phase (milli sec) |
| 5. | Speed (m/s or cm/s) | Cadence(steps/min) |
| 6. | Walk ratio (cm/step/min) | – |

Some more gait features [40] that can be analyzed for user recognition are gait variability and angular kinematics. Gait Variability (GV) can be defined as changes in gait parameters from one stride to the next. In a gait cycle, the coefficient of variation (CV) that is a measure of total variability can be calculated as *root mean square* (RMS) of standard deviation ($\sigma$) of the moment over stride period $t$ mean of the absolute moment of force over stride period using Equation 7.

$$CV = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\sigma^2}}{\frac{1}{n}\sum_{i=1}^{n}|X_i|} \tag{7}$$

Angular Kinematics of joint angles refers to the kinematics analysis of angular motion [40]. Using Equation 7, angular displacement (the difference between the initial and final angular position), angular velocity (change in angular position over a period of time), and angular acceleration (change in angular velocity over a period of time).

$$\begin{aligned} Angular\ displacement\ (\Delta\theta) &= \theta_{final} - \theta_{initial} \\ angular\ velocity\ (\omega) &= \frac{d\theta}{dt} \\ angular\ acceleration\ (\alpha) &= \frac{d\omega}{dt} \end{aligned} \tag{8}$$

### 3.7. Footstep

A footstep is defined as a combination of a single left and right stride of a person. Footstep features include stride length, stride direction, timing information, acoustic and psycho-acoustic parameters, spatial positions, and relative pressure values in foot regions. These features can be captured using a range of sensors including floor-based sensors[84], such as piezoelectric sensors, switch sensors, or fabric-based pressure mapping sensors.

Ground Reaction Force (GRF) is the common feature providing a description of a person's footstep force acquired from pressure sensors [44]. Ground Reaction Force ($GRF_i$) per sensor can be computed by accumulating each $i^{th}$ sensor pressure amplitude from time $t = 1$ to $t = T_{max}$ using Equation 9.

$$GRF_i = \sum_{t=1}^{T_{max}} P_i[t] \tag{9}$$

Furthermore, using Equation 10 time-series arrays, namely, average spatial pressure ($SP_{ave}$), cumulative spatial pressure ($SP_{cumulative}$), upper ($SP_{upper}$) and lower ($SP_{lower}$) contours can be generated from the pressure signals acquired from $N$ sensors for a $T$ time-period [85].

$$SP_{ave}[t] = \sum_{i=1}^{N} P_i[t] \qquad SP_{cumulative}[t] = \sum_{i=1}^{N} P_i[t] + \sum_{i=1}^{N} P_i[t-1]$$
$$SP_{upper}[t] = \max_{i=1}^{N} S_i[t] \qquad SP_{lower}[t] = \min_{i=1}^{N} S_i[t] \tag{10}$$

where, $P_i[t]$ is the differential pressure value from the $i^{th}$ sensors at the time $t$, and, $N$ is the total number of sensors. Footstep analysis is applicable for numerous applications, such as predicting human action, security, and surveillance at public places [85].

## 4. State-of-the-art in HCI and natural habits based behavioral biometrics

This section discusses the state-of-the-art for user recognition schemes based on HCI and natural habits-based behavioral biometrics discussed in Section 3. We present a systematic narrative of the recent literature developing touch-stroke dynamics, swipe gesture, touch signature, hand micro-movements, voice-prints, gait, and footstep behavioral biometrics modalities for designing user recognition schemes targeting IoT applications.

**Touch-stroke dynamics:** User recognition methods based on *touch-stroke dynamics* can readily implemented in IoT endpoints such as smartphones, tablets, smartwatches, or other devices equipped with a touchscreen. Zheng et al. [52] utilized users' tapping behavior for user verification in a passcode-enabled smartphone. They recruited 80 subjects to explore tapping behaviors using four different factors, i.e., acceleration, pressure, size, and time. They evaluated their scheme using a one-class classifier and achieved an EER of 3.65%. Further, their experiment to quantitatively measure the effect of the mimic attack revealed that only dissimilarity scores of acceleration reduced, whereas the score ranges of the other three features spread wider. Similarly, Teh et al. [53] investigated touch dynamics biometrics by extracting a basic set of timing and spatial features known as First Order Features (FOF). They derived an extended Set of Features (SOF) from the FOF features. They used both a one-class classifier (K-Nearest Neighbor (kNN), Support Vector Data Description (SVDD)), and a binary-class classifier (kNN, State Vector Machine (SVM)) for evaluation of their scheme on a dataset having 150 subjects. Through experiments, they demonstrated a reduction in impersonation attempts to 9.9% from 100% by integrating the touch dynamics authentication method into a 4-digit PIN-based authentication method in contrast to the sole use of PIN-based authentication.

Draw-a-pin is a PIN content analyzer and drawing behavior analyzer to verify the two factors of a log-in attempt [86]. The system extracts touch information, such as x-coordinates, y-coordinates, finger pressure, and touch area size, from each 4-digit pin. They claim the scheme is resilient against shoulder surfing attacks and achieved an EER of 4.84% using the Dynamic Time Warping (DTW) algorithm on 20 subjects. Similar to the draw-a-pin approach, Tolosana et al. [87] suggested replacing conventional authentication systems based on PIN and One-Time Passwords (OTP) with a scheme that allows users to draw each digit of the password on the device's touchscreen. They created an e-BioDigit database consisting of 93 subjects to conduct their experiment. The authors evaluated the scheme using DTW by combining with the Sequential Forward Feature Selection (SFFS) function selection algorithm and Recurrent Neural Networks (RNNs) deep learning technology that exploited various touch features; they achieved an EER of 4%.

Multi-touch authentication with TFST (*touch with fingers straight and together*) gestures is a simple and reliable authentication scheme for devices equipped with multi-touch screens [57]. The scheme exploits both hand geometry and behavioral characteristics and the authors collected a large multi-touch dataset from 161 subjects. They achieved an EER of 5.48% (5 training samples) using one-class SVM and kNN classifiers. Furthermore, they performed a security analysis for a zero-effort attack, smudge attack, shoulder surfing attack, and statistical attack. Touch-stroke dynamics is a relatively recent behavioral biometrics when compared to well established behavioral biometrics such as signature verification. Table 4 compares user recognition schemes based on touch-strokes dynamics.

Table 4: User recognition schemes based on touch-strokes dynamics

| Study | Methodology/Features | Algorithm/Classifier | Dataset | Performance |
|---|---|---|---|---|
| Li et al. [88], 2021 | Single touch, touch movement and multi-touch | SVM | 60 subjects | Average error rate ≈ 2.9% |
| Teh et al. [53], 2019 | FOF and SOF | kNN, SVDD, and SVM | 150 subjects | Impersonation rate = 9.9% |
| Zheng et al. [52], 2014 | Tapping behaviors | one-class machine learning technique | 80 subjects | EER = 3.65% |
| Song at al. [57], 2017 | Multi-touch with TFST | One-class SVM and kNN | 161 subjects | EER = 5.48% (5 training samples) |
| Tolosana et al. [87], 2017 | Handwritten numerical digits using finger-touch | DTW combined with the SFFS and RNNs | e-BioDigit [89] (93 subjects) | EER = 4% |

**Swipe gesture:** A *swipe gesture* (collection of touch-strokes from a touch-down to touch-release) can be processed for user recognition. *SwipeVlock* authenticates users based on their way of swiping the phone screen with a background image [60]. The scheme was evaluated using a decision tree, Naive Bayes (NB), SVM, and Back Propagation Neural Network (BPNN) on 150 subjects and achieved a success rate of 98%. DRIVERAUTH collected and encoded a sequence of touch-events when a user swipes on the touchscreen using their finger. It achieved a TAR of 87% using Quadratic SVM (Q-SVM) on a dataset of 86 subjects. Jain et al. [56] analyzed swipe gestures, such as left-to-right swipe (L2R), right-to-left swipe (R2L), scroll up (SU), scroll down (SD), zoom in (ZI), zoom out (ZO) and single tap (ST), subsequently, extracting x–y coordinates, accelerometer, orientation sensor readings, and area covered by a finger to design an authentication scheme. The scheme recruited 104 subjects for evaluation and 30 subjects for performance verification. Using a modified Hausdorff distance (MHD), they achieved an EER of 0.31% for combined gestures using score level fusion.

Ellavarason et al. [59] proposed a swipe gesture authentication and collected a dataset under four scenarios, i.e., sitting (room and bus) and walking (outdoor and treadmill). They used SVM, kNN, and NB are used to evaluate the robustness of swipe gestures and achieved an ERR of 1% (sitting in a room), 30% (sitting in a bus), 23% (walking on a treadmill), 27% (walking outdoor) on 50 subjects. According to Poze et al. [90], horizontal strokes hold more user-specific information and are more discriminating than vertical strokes. They investigated a statistical approach based on adapted Gaussian Mixture Models (GMM) for swipe gestures and achieved an EER of 20% (40 training samples) using a dataset with 90 subjects. Garbuz et al. [91] proposed an approach that analyzed both swipes and taps to provide continuous authentication. The one-class classification model is generated using one-class SVM. The scheme can detect an impostor in 2-3 gestures, whereas the legitimate user is blocked on average after 115-116 gestures.

Another scheme involved the extraction of temporal information from consecutive touch-strokes [92]. For evaluation, they temporal Regression Forest (TRF) architecture and achieved an EER of 4%, 2.5% on the Serwadda and Frank datasets, having 190 and 41 subjects, respectively. Kumar et al. [93] proposed a multimodal scheme that exploited swiping gestures, typing behavior, phone movement patterns while typing/swiping, and their possible fusion at the feature- and score-level for authenticating smartphone users, continuously. A multi-template classification framework (MTCF) is implemented for evaluation. They achieved an accuracy of 93.33% and 89.31% using feature level and score level fusion, respectively on 28 subjects. Table 5 compares user recognition schemes based on swipe gesture.

Table 5: User recognition schemes based on swipe

| Study | Methodology/Features | Algorithm/Classifier | Dataset | Performance |
|---|---|---|---|---|
| Jain et al. [56], 2021 | Touchscreen gestures (L2R, R2L, SU, SD, ZI, ZO, and ST) | Modified MHD | 104 subjects for evaluation and 30 subjects for performance verification | EER = 0.31% for combined gestures using score level fusion |
| Gupta et al. [58], 2019 | Touch-events sequence | Q-SVM | 86 subjects [94] | TAR = 87% |

| Study | Methodology/Features | Algorithm/Classifier | Dataset | Performance |
|---|---|---|---|---|
| Ellavarason et al. [59], 2020 | Swipe gesture in four scenarios - sitting (room and bus) and walking (outdoor and treadmill) | SVM, kNN, and NB | 50 subjects | ERR = 1% (sitting in room), 30 %(sitting in bus), 23% (walking on treadmill), 27% (walking outdoor) |
| Li et al. [60], 2020 | Swipe on an image | Decision tree, NB, SVM, and BPNN | 150 subjects | Success Rate = 98% |
| Pozo et al. [90]. 2017 | Horizontal and vertical strokes | GMM | 190 subjects | EER = 20% (40 training samples) |
| Kumar et al. [93], 2016 | Swipe, typing behavior, phone movement patterns | MTCF | 28 subjects | Accuracy = 93.33% (feature level fusion), 89.31% (score level fusion) |
| Ooi et al. [92], 2019 | Touch-strokes temporal information | TRF | Serwadda (190 subjects), Frank [95] (41 subjects) | EER = 4%, 2.5% |

**Touch-signature:** *Touch-signature* using a finger or stylus on a touchscreen device is emerging as an alternative to an all-time acceptable handwritten signature for user recognition. Features explained in Section 3.3 can be exploited to identify a user for a number of security-sensitive applications, such as hotel bookings, online-banking, and shopping thereby helping minimize fraudulent activities.

Tolosana et al. [64] proposed an on-line signature verification system that is adaptable to the signature complexity level. In their proposed approach, a signature complexity detector based on the number of lognormals from the Sigma LogNormal writing generation model, and a time function extraction module are generated for each complexity level. Then, the DTW algorithm is used to compute the similarity between the time functions from the input signature and training signatures of the claimed user. The scheme achieved an EER of 2.5% and 5.6% on BiosecurID (pen scenario of 400 subjects) and BioSign (pen and finger scenario of 65 subjects) datasets, respectively. Yoshida et al. [65] analyzed touch-strokes duration and segments' directions of signatures using two Japanese characters. An objective measure of the difference between two sequences of touching duration is used to evaluate the similarity and the scheme achieved an EER of 7.1% using 10 subjects. Gomez et al. [96] proposed to improve the performance of online signature verification systems based on the Kinematic Theory of rapid human movements and its associated Sigma LogNormal model. The authors used the BiosecurID multimodal database of 400 subjects having 6,400 genuine signatures and 4,800 skilled forgeries for the evaluation of their schemes using DTW.

Ren et al. [97] proposed a signature verification system leveraging a multi-touch screen for mobile transactions by extracting critical segments to capture a user's intrinsic signing behavior for accurate signature verification. They applied DTW to calculate an optimal match between two temporal sequences with different lengths, and then measure the similarity between them. On 25 subjects, an EER of 2%, 1%, and 3% for single-finger, two-finger, and under the observation and imitation attack scenarios, respectively achieved. Al-Jarrah et al. [98] proposed anomaly detectors, such as STD

Z-Score Anomaly Detector, Average Absolute Deviation (AAD) Anomaly Detector, and Median Absolute Deviation (MAD) Anomaly Detector, for signature verification. Using distance functions for evaluation, they achieved an EER between 3.21% to 5.44% for skilled forgeries and 4.74% to 6.31% for random forgeries among 55 subjects. Behera et al. [99] proposed an approach based on spot signature within a continuous air writing captured through Leap motion depth sensors. The processed signatures are represented using convex hull vertices and DTW is selected for performance verification of the spotted signatures. The authors achieved an accuracy of 80% on 20 subjects. Ramachandra et al. [100] proposed user verification using a smartwatch-based writing pattern or style that exploited accelerometer data acquired from 30 participants. The accelerometer data is further transformed using 2D Continuous Wavelet Transform (CWT) and deep features extracted using the pre-trained ResNet50. Table 6 compares user recognition schemes based on touch signature.

Table 6: User recognition schemes based on touch signature

| Study | Methodology/Features | Algorithm/Classifier | Dataset | Performance |
|---|---|---|---|---|
| Tolosana et al. [64], 2020 | Time functions for different complexity, Lognormals from Sigma LogNormal | DTW | BiosecurID (pen scenario of 400 subjects), BioSign (pen and finger scenario of 65 subjects) | EER = 2.5%, 5.6% |
| Al et al. [98], 2019 | finger-drawn signature | Distance-based functions | 55 subjects | EER = 3.21% to 5.44% (Skilled Forgery), 4.74% to 6.31% (Random Forgery) |
| Van et al. [86], 2017 | Touch information from 4-digit pin drawing | DTW | 20 subjects | EER = 4.84% |
| Yoshida et al. [65], 2017 | Signatures touch-strokes duration and segments directions | Distance-based | 10 Subjects | EER = 7.1% |
| Behera et al. [99], 2017 | Spot signature using leap motion | DTW | 20 subjects | Accuracy = 80% |
| Ren et al. [97], 2019 | Signature using multi-touch screen | DTW | 25 subjects | EER = 2% (for single-finger scenarios), 1% (for two-finger scenarios), 3% (under the observe and imitate attack scenarios) |

**Hand-movement:** IoT end-points equipped with motion sensors are capable of acquiring *micro-movement* produced as a result of a user's unique gesture to perform certain activities. Subsequently, the raw data collected from various sensors for an activity can be exploited when designing a user recognition scheme. SmartHandle utilizes the user's hand-movement in 3-dimensional space by determining the X, Y, and Z coordinates corresponding to the hand-movement trajectory, to generate a user-identification signature [41]. The classification model is evaluated using 3 different classifiers, i.e., the linear discriminant classifier (LDC), uncorrelated normal based quadratic Bayes classifier (UDC), and random forest (RF). The scheme achieved an accuracy of 87.27% on a dataset containing 11 subjects. Centeno et al. [101] designed an approach that acquires user-specific motion patterns using an accelerometer as a result of the user's interaction with a smartphone. The feature extraction process is based on autoencoders (a deep learning technique). On a dataset of 120 subjects, the scheme achieved an EER of 2.2%.

DeepAuth leverages time and frequency domain features extracted from motion sensors and a long short-term memory (LSTM) model with negative sampling to build a re-authentication framework using 47 subjects [102]. The authors also compared DeepAuth with state-of-the-art classification methods such as SVM, RF, Logistic Regression (LR), and Gradient Boosting (GB) classifiers and achieved an accuracy of 96.70% for the data collected for 20 seconds. Another bimodal scheme exploited touch-tapping and hands-movements while users enter the 8-digit free-text secret [54]. For the evaluation, NB, NeuralNet (NN), and RF classifiers are used and a TAR of 85.77% is achieved on 97 subjects. VeriNET employed motion signals as a password and leveraged a deep-RNN to authenticate users [103]. The scheme is evaluated on a dataset containing 310 subjects to achieve an EER of 7.17% for PINs and 6.09% for Android locking patterns.

SnapAuth profiles a user's arm-movements when the user performs a snap-action wearing smart watches [104]. The scheme was evaluated using Bayes Net (BN), Multilayer Perceptron (MLP), and RF classifiers on a dataset of 11 subjects and achieved a TAR 82.34%. Li et al. [105] proposed a continuous authentication scheme based on free-text keystroke that exploited both keystroke latency patterns and wrist motion behaviors acquired by wrist-worn smartwatches. A Dynamic Trust Model (DTM) is developed to fuse two one-vs-all RF ensemble classifiers and achieved a TAR of 98.12% on 25 subjects. Another continuous authentication scheme compares the wristband's motion with the phone's motion of a user to produce a score indicating its confidence that the person holding (and using) the phone is the person wearing the wristband [106]. A two-tier classification approach (using RF and NB binary classifiers) to correlate wrist motion with the touch input is deployed giving an accuracy of 96.5% tested with 38 subjects. A motion-based authentication method for smart wearable devices, MotionAuth, constructed users' identifiable signature by profiling their different natural gestures such as raising or lowering the arm [107]. They achieved an EER of 2.6% on a dataset of 30 users.

SilentSense exploited touch behavior (e.g., pressure, area, duration, position) and micro hand-movements (e.g., acceleration and rotation) [108]. SVM is employed to detect the identity of the current user according to each interacting behavior observation. On a dataset containing 100 subjects, SilentSense achieved an accuracy of 99%.

Similarly, Hand Movement, Orientation, and Grasp (HMOG) exploited both tapping and keystrokes modalities [109]. The features are extracted for hand micro-movements, grasp, and orientation patterns when a user taps or presses keys on a touchscreen. For the evaluation of the scheme, Scaled Manhattan with Fisher Score (SM-FS) Ranking, Scaled Euclidean with PCA (SE-PCA), and 1-Class SVM with Fisher Score (OCSVM-FR) Ranking is used. The scheme achieved an EER of 7.16% and 10.05% for walking and sitting postures, respectively, using a set of 100 subjects for the validation. Table 7 compares user recognition schemes based on hand-movements.

Table 7: User recognition schemes based on hand-movements

| Study | Methodology/Features | Algorithm/Classifier | Dataset | Performance |
|---|---|---|---|---|
| Centeno et al. [101], 2017 | Motion patterns using accelerometer | Autoencoders | 120 subjects | EER = 2.2% |
| Gupta et al. [41], 2019 | User's hand-movement in 3-D space | LDC, UDC, and RF. | 11 subjects | Accuracy = 87.27% |
| Bo et al. [108], 2013 | Touching behavior | SVM | 100 subjects | Accuracy = 99% |
| Amini et al.[102], 2018 | Time and frequency domain features from motion sensors and a LSTM model | SVM, RF, LR and GB | 47 subjects | Accuracy = 96.70% (20 seconds) |
| Mare et al. [106], 2019 | Compares the wristband's motion with the phone's motion | RF and NB | 38 subjects | Accuracy = 96.5% |
| Li et al. [105], 2017 | Free-text keystroke | DTM | 25 subjects | TAR = 98.12% |
| Buriro et al. [104], 2018 | Arm-movements to perform snap-action | BN, MLP, and RF | 11 subjects | TAR = 82.34% |
| Lu et al. [103], 2017 | Motion signals | Deep RNN | 310 subjects | EER = 7.17% (PINs), 6.09% (Android locking patterns) |
| Buriro et al. [54], 2021 | Touch-tapping and hands-movements | NB, NN, and RF | 97 subjects | TAR = 85.77 % |
| Sitova et al. [109], 2015 | Hand movement, orientation, grasp, tap and keystroke | SM-FS, SE-PCA, and OCSVM-FC Ranking | 100 subjects . Data were for sitting and walking posture | EERs = 7.16% (walking) and 10.05% (sitting) |

**Voice:** *Voice* is an easily collectible behavioral biometric modality that can be acquired by any IoT end-point equipped with a microphone. Section 3.5 has explained the features that are normally exploited for designing voice-based user recognition schemes.

An automatic voice biometric authentication scheme that recognizes a speaker using MFCC and Discrete Cosine Transform (DCT) is presented in [110]. On a dataset of 13 subjects, a SVM using radial-basis function (RBF) kernel is used for evaluation, achieving a success rate of 90%. DriverAuth computed statistical features after extracting MFCCs from a bandpass filter voice signal containing 2 channels sampled at 44,100 Hz with 16 bits per sample [58]. The authors used Q-SVM, ETB, Weighted kNN (W-kNN) classifiers for generating a multi-class classification model. On a dataset of 86 subjects, the system achieved a TAR of 90.5% with voice features and 95.1%

with voice and swipe features combined.

Doddappagol et al. [111] proposed text prompted voice recognition system that used MFCCs, Pitch and Formant technique for extracting features. On a dataset containing 25 subjects, with SVM employed for user classification, an accuracy between 88.7% and 92% was achieved. BreathPrint exploits the audio signatures, i.e., sniff, normal, and deep breathing, of a person [112]. A microphone sensor in close proximity to users' nose acquires these three audio signatures produced by them. A classification pipeline using Gammatone Frequency Cepstral Coefficients (GFCC) as features as part of a GMM based classifier was used for evaluation, and achieved an accuracy of 94% on a dataset comprising 10 subjects. VoiceLive performs liveness detection by measuring Time-Difference-of-Arrival (TDoA) changes for a sequence of phoneme sounds [68]. It evaluates a phoneme sound localization based liveness detection system that distinguishes a passphrase spoken by a live user from a replayed one giving an accuracy of 99% on a dataset containing 12 subjects. Table 8 compares user recognition schemes based on voice-print.

Table 8: User recognition schemes based on voice

| **Study** | **Methodology/Features** | **Algorithm/Classifier** | **Dataset** | **Performance** |
|---|---|---|---|---|
| Doddappago et al. [111], 2016 | MFCCs, Pitch and Formant technique | SVM | 25 subjects | Accuracy = 88.7% to 92% |
| Chauhan et al. [112], 2017 | Audio signatures (sniff, normal, and deep breathing) | A GFCC and GMM | 10 subjects | Accuracy = 94% |
| Zhang et al. [68], 2016 | Spoken passphrase | Liveness detection by measuring TDoA changes for a sequence of phoneme sounds | 12 subjects | Accuracy = 99% |
| Barbosa et al. [110], 2015 | MFCC and DCT of voiceprint. | SVM-RBF | 13 subjects | Success Rate = 90% |
| Gupta et al. [58], 2019 | Statistical features from MFCCs | Q-SVM. | 86 users | TAR = 90.5% |

**Gait:** The *human gait* is a spatio-temporal motor-controlled biometric behavior that can be employed for to recognise individuals unobtrusively, using a camera, radar, position-, motion-, or pressure-based sensors. Musale et al. [113] proposed a Lightweight Gait Authentication Technique (Li-GAT) that exploits information, such as the subconscious level of user activities, collected from IoT devices having inbuilt motion sensors including an accelerometer. For evaluation, LR using deep-NN, RF, kNN classifiers were selected and achieved an accuracy of 96.69% on a dataset containing 12 subjects. Kastaniotis et al. [114] designed a gait recognition system based on a hierarchical representation of gait trajectories acquired using depth and motion sensors. The acquired pose sequences are expressed as angular vectors (Euler angles) of eight selected limbs. These trajectories (sequences of angular vectors) are then mapped in the dissimilarity space, resulting in a vector of dissimilarities that are modeled via sparse representation. For verification, three criteria were evaluated: the

Sparsity Concentration Index (SCI), the minimum dissimilarity (MinDiss), and the combination of both, and achieved an EER of 3.1% on 30 subjects.

Deep Gait authenticates users based on a single walk cycle [115]. It acquires accelerometer and gyroscope readings from wearable or hand-held devices to determine a users' gait. For evaluation, a deep-NN is used that achieved an EER of 1.8% on 51 subjects. Another smartphone-based gait recognition system with the application of Subjective Logic (SL) for biometric data fusion is presented in [116]. Gait features considered for the system are statistical (ST), the histogram of the distribution (BIN), MFCCs, and Bark-frequency cepstral coefficients (BF1 and BF2). For evaluation, Extremely Randomized Trees (ERT), MLP, and RF classifiers are selected that gave an EER of 1.31% on 48 subjects. Lamiche et al. [117] proposed a bimodal authentication scheme based on gait patterns and keystroke dynamics. By using the smartphone's built-in sensors, the user's gait signals with keystroke dynamics are acquired simultaneously, during walking and text typing activities. The scheme was evaluated using 20 subjects and an accuracy of 99.11% is achieved using a MLP classifier.

Gait-Watch is a context-aware gait-based authentication system, which is coupled with a smart-watch based activity detector to identify a user's current activity [118]. As per the real-time input of the activity detector, identification is performed on corresponding training templates. The method extracted unique features of gait dynamics by exploiting the scale-space of gait acceleration signals using a sparse coding scheme. For identification, probabilistic sparse representation classification (PSRC) is employed and the method achieved 97.3% recognition accuracy and 3.5% EER. An improvement of 30.21% in recognition accuracy is observed by dynamically determining the user's activity. Table 9 compares user recognition purposes based on a user's gait.

Table 9: User recognition schemes based on gait

| **Study** | **Methodology/Features** | **Algorithm/Classifier** | **Dataset** | **Performance** |
|---|---|---|---|---|
| Wasnik et al. [116], 2017 | Users' gait ST, BIN, MFCCs, BF1 and BF2 | ERT, MLP and RF | 48 subjects | EER = 1.31% |
| Musale et al. [113], 2018 | Walking based activities | deep-NN, RF, kNN | 12 subjects | Accuracy = 96.69% |
| Kastaniotis et al. [114], 2015 | Gait trajectories | SCI, MinDiss and their combination | 30 subjects | EER = 3.1% |
| Bael et al. [115], 2019 | Single walk cycle using motion sensors | deep-NN | 51 subjects | EER = 1.8% |
| Lamiche et al. [117], 2019 | Gait patterns and keystroke dynamics | MLP | 20 subjects | Accuracy = 99.11% |

**Footstep:** *Footstep features* to recognize a person can be collected imperceptibly using pressure-based sensors. Moreover, people can be allowed to walk over the footstep sensors wearing footwear (*such as shoes, trainers, boots*) and carrying weights (*such as shoulder bags and files*) that make the recognition process more realistic.

Rodriguez et al. [119] proposed a scheme that exploits footstep signals in both the time and space domains. In the time domain, the extracted features include the ground reaction force (GRF), the spatial average, and the upper and lower contours of the pressure signals; the spatial domain, involves features including $3D$ images of the accumulated pressure. A SVM-RBF is used for evaluation. On a dataset of 120 subjects, EERs of 15.2%, 13.4%, and 7.9% were achieved, by a training classification model with 40, 100, and 500 single footstep signals respectively, after fusing both time-domain and space-domain features. Similarly, Edward et al. [44] extracted geometric and wavelet features from a footstep dataset collected by the Swansea University Speech and Image Research Group. On a dataset of 94 subjects, the scheme achieved an EER 16.3% using the RF classifier for individual prediction.

Zhou et al. [120] proposed a user identification scheme based on a single footstep biometric without considering the shape details or inter-step relationships of users' footprints. They utilized fabric sensors to register features such as shifting of the center of gravity, maximum pressure point, and overall pressured area. Evaluation of the scheme was performed using Q-SVM and it achieved an accuracy of 76.9% on a dataset containing 529 footsteps collected from 13 subjects.

One automatic biometric verification scheme leveraged spatio-temporal footstep representation acquired from floor-only sensor data [85]. For evaluation, an ensemble of a deep resnet architecture and SVM models were used and achieved an EER of 0.7% on 120 subjects. Riwurohi et al. [121] proposed a biometric identification system based on the sound of footsteps acquired using microphone arrays. The footstep sound features of 10 participants were extracted using MFCCs. The scheme achieved an accuracy of 98.8% using BPNN. Table 10 compares user recognition schemes based on a user's footstep.

Table 10: User recognition schemes based on footsteps

| Study | Methodology/Features | Algorithm/Classifier | Dataset | Performance |
|---|---|---|---|---|
| Edward et al. [44], 2014 | Extracted geometric and wavelet features from a footstep. | RF | 94 subjects | EER = 16.3% |
| Vera et al. [119], 2013 | Time and space domains footstep signals. | SVM-RBF | 120 subjects | EERs = 15.2%, 13.4%, and 7.9% with 40, 100, and 500, respectively |
| Costilla et al. [85], 2018 | Spatio-temporal footstep representations | Deep resnet architecture and SVM | 120 subjects | EER = 0.7% |
| Zhou et al. [120], 2017 | Single footstep signal with inter-step relationships | Q-SVM | 13 subjects | Accuracy = 76.9% |
| Riwurohi et al. [121], 2018 | Footsteps' sound | BPNN | 10 subjects | Accuracy = 98.8% |

## 5. Security, Privacy and Usability Considerations

*Security*, *privacy* and *usability* are indispensable non-functional requirements for designing human-to-things recognition schemes [122] that satisfy `CIA` criteria, i.e., confidentiality (ensuring access to legitimate users only), integrity (guaranteeing modification by legitimate users) and availability (ensuring uninterrupted availability to legitimate users). With regard to these requirements, substantial improvements can be observed in evolving behavioral biometric-based user recognition schemes for `AAA`.

### 5.1. Security

Reportedly, a number of security analyses have been performed to evaluate touch-based recognition mechanisms against common attacks such as impersonation, mimicking, smudge or shoulder-surfing [52, 53]. Sewadda et al. [123] rigorously analyzed the impact of *Lego-driven robotic attacks*, namely, population statistics-driven and user-tailored attack on touch-based authentication. In a population statistics-driven attack, patterns are acquired from a large database to train the robot, whereas, in a user-specific attack, samples of a legitimate user are stolen to train the robot. Subsequently, both attacks were launched by a Lego robot trained to swipe on the touch screen. Further, these attack methods can be refined for standard impostor testing for touch-based recognition schemes. Song et al. [57] conducted a security analysis of their TFST gesture authentication against: *zero-effort attack*, i.e., an adversary attacks without any prior knowledge of the underlying authentication scheme; *smudge attack*, i.e., an adversary manages to identify and trace oily residues on a touchscreen; *shoulder surfing attack*, i.e., an adversary secretly observes the legitimate user; and *statistical attack*, i.e., an adversary employs knowledge obtained from the statistics of a group of users.

A Continuous Smartphone Authentication Method using wristbands (CSAW) exploited motion gestures to verify whether a smartphone is in the hands of a legitimate owner or not [106]. Security analysis for CSAW is performed against: *opportunistic snooping*, i.e., an adversary snoops into other smartphones when the owner is not around; *stealing credentials*, i.e., an adversary steals the credentials for accessing smart devices remotely; and *shadowing*, i.e. an adversary shadows a user to access his/her smartphone illegitimately. They reported a false-positive rate of less than 2%. Yi et al. [124] performed an empirical study on the security and usability of a real-time free-form motion gesture authentication scheme (REMOTE) that leveraged user-created 3D gestures. They evaluated REMOTE against: *random attacks*, i.e. an adversary does have any prior knowledge of the victim's gesture and apply random guess to attack; *content-aware attack*, i.e., an adversary has the descriptive information about the victim's gesture obtained via social engineering or a third party; and *mimicry attack*, i.e., an adversary observes a legitimate user's gesture directly or through a recorded video. The authors reported that random attacks are ineffective for attacking gesture-based behavioral biometric authentication. In the case of content-aware attacks, additional descriptive information provides only minimal help to adversaries. Although, mimicry attacks seem more effective than the random and content-aware attacks, they still only achieve negligible success in most of the attack attempts.

Many studies have been performed to understand common attacks on voice-based recognition systems. VAuth [125] exploited users' language, accent, or mobility to ensure voice assistants - such as Siri, Google Now and Cortana - execute the commands that originate only from the voice of the owner. VAuth successfully averted attacks, such as replay-, voice-mangling, and impersonation attacks using a multi-stage matching algorithm. Rahmeni et al. [126] proposed a method to mitigate spoofing attacks, such as impersonation, replay, voice-conversion, and speech-synthesis independent of an attack-type. Their proposed method decomposes the speech signal into a glottal source signal and models the vocal tract filter using glottal inverse filtering. Features are obtained using Iterative Adaptive Inverse Filter (IAIF) descriptors that can be exploited to distinguish between genuine or spoofed input speech using a SVM and an extreme learning machine (ELM).

Chang [127] proposed a two-layer authentication method using a voiceprint to mitigate replay attacks. Similarly, the VoiceLive system addressed a replay attack using extracts of the TDoA of each phoneme sound to distinguish between a passphrase spoken by a live user and a replayed one. It leverages the human speech production system and advanced smartphone audio hardware. Garg et al. [128] investigated the effectiveness of Constant-Q Cepstral Coefficients (CQCC) and MFCC features extracted from individual frequency subbands to improve the performance of replay attack detection in automatic speaker verification (ASV) systems. Tom et al. [129] proposed group delay (GD) grams that can be obtained by concatenating a group delay function over consecutive frames as a novel time-frequency representation of an utterance. Subsequently, GD-grams provides a time-frequency representation with a high spectral resolution that can be used for the end-to-end training of deep-convolutional NNs to detect audio replay attacks.

Voice conversion attacks apply synthetic speech generation or source voice morphing to achieve the same effect as human impersonation or adapted speech synthesis, thus, deceiving the speaker identification (SID) and speaker verification (SV). An approach exploited score-level fusion of front-end features, namely, CQCCs, all-pole group delay function (APGDF), and fundamental frequency variation (FFV) to detect a synthetic speech [130]. Similarly, Yang et al. [131] investigated the high-frequency-based features for the detection of spoofing attacks. The method analyzed inverted constant-Q coefficients (ICQC) and inverted CQCC using DCT on inverted octave power spectrum and inverted linear power spectrum respectively, to detect synthetic speeches. Wu et al. [132] reported that a hidden Markov model (HMM) based text-dependent systems with temporal speech information provided more resistance to voice conversion attacks than systems lacking temporal modeling.

Munaz et al. [133] evaluated the security strength of a smartphone-based gait recognition system against zero-effort and live minimal-effort impersonation attacks, under realistic scenarios using live visual and audio feedback. Particularly, live impersonation attacks were performed by five professional actors specialized in mimicking body movements and body language. They reported no false positives under impersonation attacks and 29% of attacks were completely unsuccessful. Gait-Watch was evaluated against the imposter attack scenario [118] and reported a false acceptance of only 3.5 per 100 impostor trials. ZEMFA [134], a zero-effort multi-factor authentication system for securing access to a terminal, leveraged a smartphone and smartwatch (or bracelet) to acquire gait patterns, i.e., mid/lower body movements measured using the phone and wrist/arm movements using the watch. The scheme reported 0.2% false negatives and 0.3% false positives on average for passive attacks under benign settings. Further, the authors reported 4.55% false positives on average for active imitation attacks, such as treadmill-based attacks. Tram et al. [135] proposed a technique to prevent statistical attacks due to the inter-class low-discrimination and intra-class high-variation of gait data. The proposed technique leveraged Linear Discrimination Analysis (LDA) to enhance the discrimination of gait templates, and Gray code quantization to extract high discriminative and stable binary template that can significantly improve the security and performance of inertial-sensor based gait cryptosystem.

Moreover, behavioral biometrics have been evaluated for designing implicit [136, 137, 138], continuous [91, 93, 117], and risk-based [54, 139] user recognition schemes. Although, more comprehensive security evaluations of these behavioral biometric modalities are desired to avert any unauthorized intrusion by adversaries, repudiation claims by malicious users, denial-of-service to legitimate users, or users' privacy erosion due to function creep [140].

### 5.2. Privacy

Privacy-preserving techniques [141], such as *Template Protection Schemes*, *Biometric Crypto-Systems*, and *Pseudonymous Biometric Identities* can be implemented to safeguard users' biometric data to address issues arising from concerns in areas such as *irreversibility*, *revocability*, *unlinkability*, and *discriminability*. There are an increasing number of regional, national and international privacy protection laws and regulations, such as [142, 143, 144], that place biometric modalities under a special category of personal data. ISO24745:2011 [145] defines the following 4 properties for a template protection scheme:

- *Irreversibility*: Reconstruction of original biometric features from a stored biometric template must be computationally exhaustive to discourage adversaries to reconstruct the biometric data from features in protected form.
- *Revocability*: Ability to generate multiple versions of secure biometric templates from the same biometric data of a user that can enable the replacement of the compromised biometric template with a new template instantaneously, without causing any inconveniences to the user.
- *Unlinkability*: Multiple biometric templates of the same subject used by different recognition systems must not allow identifying/linking the user based on protected features.
- *Discriminability*: Secure template must not degrade the recognition accuracy of a biometric-based recognition system and should maintain sufficient discriminative information from rest of the registered users.

Some of the basic techniques for generating cancelable biometric templates are based on noninvertible geometric transformations, such as affine, cartesian, polar, or functional transformation [146]. Bioconvolving [147] can be useful for all the behavioral biometric modalities in which raw signals are a sequence of real-numbers of finite length. In this method, each transformed sequence can be obtained from the corresponding original sequence having $N$ values by dividing the original sequence into $W$ non-overlapping segments ($W < N$) using randomly selected $W$ integers between 1 and 99 in the ascending order. Zhi et al. [148] proposed learning-based Index-of-Maximum (LIoM) hashing that utilizes a supervised learning mechanism to generate a more discriminative and compact cancelable touch-stroke template. With a supervised learning approach, the LIoM learns the optimized projection itself, unlike data-agnostic IoM hashing that depends on random projection for hashing. The authors reported that the classification model generated with a protected template achieved significantly better accuracy than with an original template.

Chee [149] proposed Random Binary Orthogonal Matrices Projection (RBOMP) and Two-dimensional Winner-Takes-All (2DWTA) hashing for voice template protection. RBOMP transforms a 1-D voice feature (i-vector having a fixed-length real value representation) from a linear space into an ordinal space by convolving with a binary orthogonal matrix. Further, a user-specific random token and a non-invertible function such as prime factorization are used to conceal the returned index that strengthens the system security significantly. Conversely, 2DWTA hashing transforms a 2-D feature from a continuous value to a discrete value. 2DWTA relies on an implicit ordering of the feature rather than the absolute feature value of the features. That is, 2DWTA hashing defines an ordinal embedding with an associated rank-correlation measure. Billeb et al. [150] proposed a fuzzy commitment scheme by employing binarized feature vectors in a cryptographic primitive for voice features that are extracted with a speech recognition system based on GMM and UBM (Universal Background Modeling). The proposed binarization scheme generates fixed-length binary voice templates.

Elrefaei et al. [151] proposed a fuzzy commitment scheme to protect gait features extracted from gait images of one complete gait cycle using a local ternary pattern (LTP). The final feature vector is produced using principal component analysis (PCA) on the average images concatenated using a 2D joint histogram. Further, to enhance the robustness of the system, only highly robust and reliable bits from the feature vector are extracted. Bose–Chaudhuri–Hocquenghem (BCH) codes are used for key encoding and decoding during the enrolment and verification phase, respectively. Similarly, Rúa et al. [152] proposed a Hidden Markov Model-Universal Background Model (HMM-UBM) gait authentication system that incorporated template protection based on a fuzzy commitment scheme. The authentication succeeds only when the Hamming distance between the binary representation obtained during the verification and the one stored at the time of the enrollment is equal to, or less than, the error-correcting capability of the employed Error Correcting Code (ECC).

In addition, hardware-level encryption can be employed on client devices to establish trust between users and businesses as a part of a privacy-first approach for behavioral analytics. A biometric system in an IoT setting becomes unusable if it is unable to revoke biometric templates and avoid biometric template leakage as multiple services rely upon same biometric modalities from each user. Comparatively, issues related to user privacy in employing behavioral biometrics are less invasive than biological biometrics; it is strongly recommended to include an appropriate template protection scheme for designing behavioral biometric-based user authentication schemes.

### 5.3. Usability

This section discusses how behavioral biometrics for user recognition schemes can meet the guidelines defined by ISO 9241-11 standard [153]. This standard defines usability as "*the extent to which a product can be used by specified users to achieve specific goals with effectiveness, efficiency, and satisfaction in a specified context of use*". Furthermore, we describe how these attributes can be used for quantifying the usability of a user recognition system.

Still et al. [154] presented a set of human-centered authentication design guidelines. The guidelines for usable security included the need for transparent authentication process, no modality overheads on users' limited working memory, to support inclusivity, and to provide faster access. Generally, usability evaluation methods (UEM) incorporate techniques such as inspection, testing, or surveying, to assess the extent to which usability objectives are achieved for a user recognition system. The usability evaluation processes can be *formative*, i.e., evaluation performed during the design and development phase of a system, or *summative*, i.e., evaluation based on users' assessment after they use the system [155].

A number of behavioral biometric-based user recognition schemes rely on a System Usability Scale (SUS) for the subjective assessments of their usability [86, 156, 157]. VAuth conducted a usability survey using Amazon Mechanical Turk [125]. TFST gesture authentication evaluates its usability by comparing to the commonly used methods of passcode and pattern lock mechanisms [57]. They determine the usability from four different perspectives: 1) Is it easy to memorize?; 2) Is it fast to login?; 3) Is it convenient to perform authentication?; and 4) Is it less error-prone? For each question, users could respond as "disagree", "neutral" or "agree". UEMs and surveys can help to analyze perceived usability and user experiences for a user recognition scheme to ensure wider acceptance from users.

As illustrated in Figure 6, we recommend a holistic method for computing intrinsic usability attributes that can impact end-users' decision to use a security mechanism: *effectiveness*, *efficiency*, *satisfaction*, *thoroughness*, *validity* and *reliability*. Equations 11 to 16 can be applied to measure usability attributes empirically, for a user recognition scheme by employing a UEM.

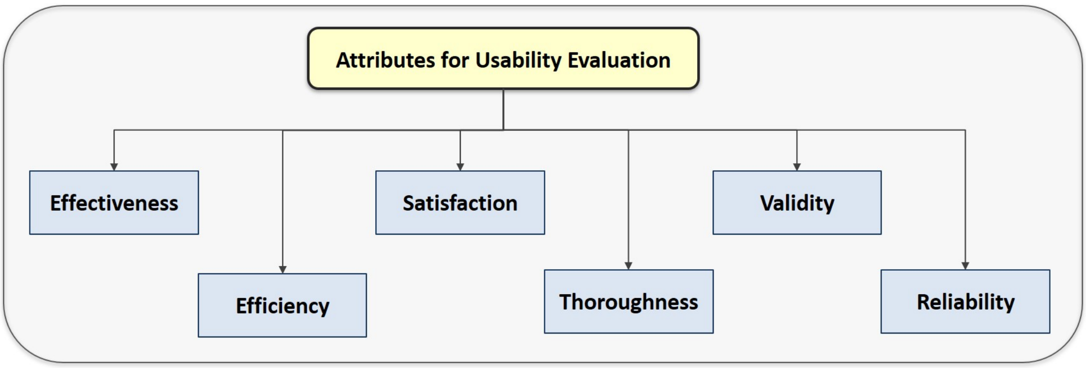


Figure 6: Attributes for usability evaluation

Effectiveness [158] is the degree to which users correctly and completely achieve specified goals and it can be measured using Equation 11.

$$Effectiveness = \frac{Goals\ achieved\ successfully}{Total\ number\ of\ goals} \times 100\% \tag{11}$$

Efficiency [158] can be measured using speed and interactiveness using Equation 12.

$$\begin{aligned} Efficiency_{speed} &= StopTime_{milliseconds} - StartTime_{milliseconds} \\ Efficiency_{interactiveness} &= Count(Number\ of\ Steps) \end{aligned} \tag{12}$$

Satisfaction [158] can be measured using Equation 13, which is an average of all the responses to a post-task questionnaire questions. Questionnaire responses can be an ordinal value, e.g., Linkert scale (1 = Strongly disagree to 5 = Strongly agree).

$$Satisfaction = \frac{\sum_{n=1}^{N} Response_n}{N} \tag{13}$$

Thoroughness [159] of a user recognition scheme concerning all of the identified usability issues can be measured using Equation 14. A UEM is expected to determine all the possible usability issues with respect to a user recognition scheme.

$$Thoroughness = \frac{Number\ of\ real\ usability\ issues\ identified}{Number\ of\ real\ usability\ issues\ exist} \tag{14}$$

Validity [159] to assert the correctness of the UEM results can be measured using Equation 15.

$$Validity = \frac{Number\ of\ real\ usability\ issues\ identified}{Number\ of\ all\ usability\ issues\ identified} \tag{15}$$

Reliability [159] to determine the consistency of a UEM, regardless of the individual performing the usability evaluation, can be measured using Equation 16.

$$Reliability = \frac{Number\ of\ usability\ issues\ identified\ by\ each\ user}{Number\ of\ usability\ issues\ identified\ by\ at\ least\ one\ user} \tag{16}$$

During the design phase of a user authentication scheme, UEMs can effectively embody these attributes to indicate the overall usability. A relationship between the system architecture and given sets of usability requirements can be derived using Equations 11 - 16. This enables both software engineers and usability specialists to evaluate whether the system is ultimately usable. These metrics enable usability specialists to determine which aspects of usability require redress. Subsequently, software engineers can evaluate how these aspects of usability can be fulfilled within the context of the architecture without affecting vital quality attributes, such as security, performance, availability, time and cost. Usability is a significant quality attribute, or non-functional requirement, since in cases that the human-to-things recognition scheme is unusable, users will either compromise the function to make it more usable, or avoid using completely.

### 5.4. User Recognition Scheme Readiness

While designing a user authentication scheme, the attributes - *security*, *privacy*, and *usability* are often perceived as orthogonal to each other. Studies have shown that available user recognition schemes struggle to satisfy these three attributes simultaneously [160]. We introduce a dashboard that is a $2 \times 2$ matrix having usability and privacy status indicators as rows and columns to interpret a user recognition scheme readiness, as illustrated in Figure 7.

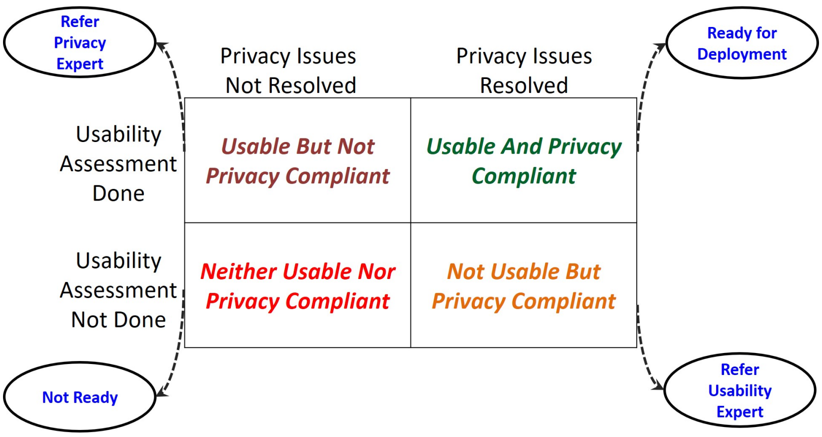


Figure 7: A dashboard for a user recognition scheme readiness

The dashboard can be useful when the user recognition scheme is baselined after incorporating a given set of security requirements. User recognition scheme qualifying to the Top-Right block of the dashboard indicates the scheme is usable and privacy-compliant, i.e., ready for deployment. Section 5.2 can be referred if the scheme qualifies to the Top-Left block, i.e., usable but not privacy-compliant. Section 5.3 can be referred if the scheme qualifies to the Bottom-Right block, i.e., not usable but privacy-compliant. The scheme is not ready if it only qualifies to the Bottom-Left block, i.e., neither usable nor privacy-compliant.

## 6. Open Challenges and Opportunities

This section presents the limitations of current approaches to designing behavioral biometric-based authentication schemes and outstanding challenges followed by general prospects and opportunities. It is worth emphasizing that HCI and natural habit-based behavioral biometrics have the power to reshape the human-to-things recognition market in the next few years.

### 6.1. Challenges and Limitations

Given the heterogeneity of behavioral biometric modalities, the limitations and vulnerabilities associated with each modality must be investigated during the conceptualization phase of a behavioral biometric-based user recognition system.

- Recently, deep generative models (DGMs) such as Generative Adversarial Networks (GANs) or Variational Autoencoders (VAE) have been adopted to generate attacks on biometric-based recognition systems and these represent a significant emerging challenge [161]. A thorough testing strategy for liveness-detection, intra-class variance and common attacks (e.g., malware, mimic, impersonation, spoofing, replay, statistical, algorithmic, and robotics attack) mitigation [29] must be developed as part of the security analysis.

- Privacy regulation laws, such as General Data Protection Regulation (GDPR) [142], the California Consumer Privacy Act (CCPA) [143] and the Health Insurance Portability and Accountability Act (HIPAA) [144], mandate an increase in responsibility and transparency for using and storing personal data. According to GDPR, biometric data that allow or confirm the unique identification of an individual is recognized as a special category of personal data under Art. 9 [162]. Consequently, there is a need to employ adequate measures (e.g., template protection and template storage location) for users' privacy conformance as per these laws.

- Another important aspect that requires addressing concerns the ethical risks in the use of behavioral biometrics [163]. Recording of data for behavioral biometric modalities over time could result in the dynamic behavior profiling of a person, which can reveal how the person has behaved in a certain context. Particularly, this can become more critical when modalities are combined with soft biometrics, such as age, gender, height, weight and ethnicity, since this can generate a more sensitive profile of a person. The creation of sensitive profiles can lead to ethical risks, such as: *discrimination* - for example to exclude a person from certain areas and activities; *stigmatization* - to create a negative interpretation of a person; and *unwanted confrontation* - the disclosure of personal information (for example, body signals may indicate a certain disease or cognitive ability of a person).

- Quality control of the biometric template is a prerequisite before the enrollment or verification/identification step [164]. This can support the correctness, consistency, redundancy and speed of a biometric system to overcome problems arising from the sensors, environment or users themselves.

- Certain factors such as aging, fatigue, stress, mood, sleep deprivation, injury and disease could inhibit the effectiveness of behavioral biometric modalities. These factors also require a thorough investigation to support the evolution of behavioral biometric-based recognition systems.

- Behavioral biometrics datasets are required to include all the demographics, such as covering different age groups, cultural factors and ethnicity, to provide better objectivity. Further, standards for behavioral biometrics and benchmarking of sensors must be developed and utilised.

### 6.2. Prospects and Opportunities

Behavioral biometrics have the potential to deliver secure, transparent, continuous and cost-effective human-to-things recognition solutions for emerging IoT ecosystems. They can offer multi-faceted benefits: 1) behavioral biometric modalities can be collected transparently (non-intrusive) [165]; 2) the availability of a wide range of sensors *(e.g., Accelerometer, Gyroscope, Radar, Piezometer, Microphone and Proximity sensors)* enable acquisition of behavioral biometric modalities accurately and efficiently; 3) they can be leveraged for designing implicit (frictionless) [136], continuous (active) [33, 42] or risk-based (non-static) recognition systems due to the evolution of embedded Machine Learning engines [166]; 4) they do not add cognitive load on users; 5) they cannot be easily stolen, shared, transferred, conjectured or hacked; and 6) they are, comparatively, less prone to cyber-attacks [122].

Sensors to capture behavioral biometric modalities are advancing rapidly, both in scope and technology. With the emergence of fabrication techniques such as Micro-Electro-Mechanical Systems (MEMS), microminiaturized sensors, actuators, mechanical components and electronics can be integrated into a single chip [167]. ST Microelectronics is one of the leading MEMS manufacturers that provides high-performance sensors with ultra-low power requirement [168]. RoKiX Sensor Node integrates multiple sensors with Bluetooth Low Energy (BLE) interface to provide the measurement of 3D-acceleration, 3D-magnetism, 3D-rotation, pressure, and temperature [169]. A wide range of touch screen (such as 5-Wire Resistive, Surface Capacitive touch, Projected Capacitive (P-Cap), Surface Acoustic Wave (SAW) and Infrared (IR) [170]) sensors are available in the market that can be selected for ATMs, kiosks, vending machines, smart devices or wearables' screens. High-performance piezoresistivity, capacitance or piezoelectric pressure sensors can be miniaturized using silicon fabrication techniques, for example piezoelectric based insole sensor [171]. Time-of-Flight 3D sensors utilise Light Detection and Ranging (LIDAR) to measure distances and sizes, to track motions, and to convert the shape of objects into 3D models [172, 173].

Operating systems such as *Android*, *iOS*, *Windows* provide SDK and APIs for interfacing sensors to acquire behavioral biometric modalities [174, 175, 176]. Leading system on a chip (SoC) manufacturers and designers, such as Intel and ARM provide SoCs supporting machine learning engines [23], AI-embedded chips [177] and NN-powered FPGAs [178] capable of supporting advanced algorithms for sensor data fusion, learning autonomously from existing data, acquiring knowledge for assessments, and making predictions and decisions. Further, IoT platforms, such as Google Cloud, IBM Watson, Amazon AWS, Microsoft Azure support advanced machine learning, and Artificial Intelligence algorithms backed by enormous computational power that can provide the necessary infrastructure to design behavioral biometric-based user recognition systems for a variety of applications. Thus these advances will continue to deliver further enhanced capabilities for behavioral biometric-based user recognition.

Key market players, particularly, *BehavioSec*, *BioCatch*, *EZMCOM*, *NEC Corporation*, *SecuredTouch* have been exploiting behavioral biometrics to design security solutions for financial institutions, businesses, government facilities, e-commerce merchants and online businesses to support security-sensitive applications. The security solutions offered range from prevention of the use of stolen or synthetic identities in applying for credit online to making better fraud decisions. Solutions can be deployed as an extra layer of intelligence to support user recognition in the fight against cyber-crimes.

Table 11: IoT domains, key applications and behavioral biometrics usage

| **IoT Domains** | **Key Applications** | **Touch-stroke** | **Swipe** | **Touch Signature** | **Hold movements** | **Voice** | **Gait** | **Footstep** |
|---|---|---|---|---|---|---|---|---|
| Smart infrastructure | Smart homes, smart offices, smart cities, smart grid, Waste management, social networking apps | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Transportation | Smart ticket booking, intelligent access system, smart parking, driverless Taxis | ✓ | ✓ | ✓ | | ✓ | | |
| Healthcare | Smart hospital, medical records | ✓ | | ✓ | | ✓ | | |
| Industrial control | Smart retail, supply chain management | | | ✓ | | ✓ | | |
| Security surveillance | Perimeter access control, border control, intrusion detection systems | ✓ | | | ✓ | | ✓ | ✓ |

Behavioral biometrics can offer opportunities to address the security and usability issues that end-users can face when using conventional user recognition schemes. Table 11 suggests IoT domains, key applications, and behavioral biometrics that can be exploited for user recognition. If not replacing conventional mechanisms entirely, behavioral biometrics can minimize the burden placed on them to security-sensitive IoT ecosystems [166]. Another benefit of behavioral biometrics is that they can be fused with each other, and with biological biometrics, seamlessly to build more robust recognition schemes. Security-sensitive sectors such as smart banking, e-commerce and finance are already leveraging behavioral biometric-based user recognition mechanisms [165]. Furthermore, HCI-based behavioral biometrics can be applied to minimise cyber-abuse and online scams, such as the spread of fake news, creation of bogus profiles on social media platforms, phishing, as well as similar illegal activities.

## 7. Conclusions

Within the overall IoT security spectrum, robust and usable *human-to-things* recognition schemes are of increasing importance, given the highly prescriptive nature of conventional (knowledge- or token-based) recognition schemes currently being utilised. The efficacy of conventional schemes remains limited since they require users to recall something they know or to possess something. As such, user recognition schemes for emerging IoT ecosystems, which can fulfill both the security and usability criteria, and comply the privacy laws, are in genuine demand.

This article has summarized the state-of-the-art in HCI- and natural habits-based biometrics, namely, touch-stroke, swipe, touch-signature, hand-movements, voice, gait and footstep. Attributes and features for each of these identified and analysed so that they can be best exploited in the design of user-friendly recognition schemes. A discussion of security, privacy and usability evaluation indicators together with the existing challenges and limitations is also presented that requiring attention to achieve the widespread adoption of behavioral biometric-based recognition schemes.

Overall, the prospects and market trends cited in this article indicate that behavioral biometrics can provide innovative ways to implement implicit (*frictionless*), continuous (*active*) or risk-based (*non-static*) recognition schemes. With the availability of smart sensors, advanced machine learning algorithms and powerful IoT platforms, behavioral biometrics can replace conventional recognition schemes, thereby reshaping the existing user recognition landscape for IoT ecosystems.

## Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation programme DS 2018-2019-2020 as part of the E-Corridor project (www.e-corridor.eu) under grant agreement No 883135. Professor Maple would like to acknowledge the support of UKRI through the grants EP/R007195/1 (Academic Centre of Excellence in Cyber Security Research - University of Warwick), EP/N510129/1 (The Alan Turing Institute) and EP/S035362/1 (PETRAS, the National Centre of Excellence for IoT Systems Cybersecurity). This work was also supported, in part, by the Bill & Melinda Gates Foundation [INV- 001309]. Under the grant conditions of the Foundation, a Creative Commons Attribution 4.0 Generic License has already been assigned to the Author Accepted Manuscript.

## References

[1] Harvard, Technology factsheet series: Internet of things, https://www.belfercenter.org/sites/default/files/2019-06/TechFactSheet/iot%20-%205.pdf, online web resource (*Accessed on 30-06-2021*).

[2] B. Bera, A. K. Das, W. Balzano, C. M. Medaglia, On the design of biometric-based user authentication protocol in smart city environment, Pattern Recognition Letters 138 (2020) 439–446.

[3] S. N. Swamy, D. Jadhav, N. Kulkarni, Security threats in the application layer in iot applications, in: Proceedings of the International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), IEEE, 2017, pp. 477–480.

[4] M. Trnka, T. Cerny, N. Stickney, Survey of authentication and authorization for the internet of things, Security and Communication Networks 2018.

[5] Verizon, Data breach investigations report, https://enterprise.verizon.com/resources/reports/dbir/, online web resource (*Accessed on 30-06-2021*).

[6] E. Fernandes, A. Rahmati, K. Eykholt, A. Prakash, Internet of things security research: A rehash of old ideas or new intellectual challenges?, IEEE Security & Privacy 15 (4) (2017) 79–84.

[7] H. Lin, N. W. Bergmann, Iot privacy and security challenges for smart home environments, Information 7 (3) (2016) 44.

[8] A. Ometov, V. Petrov, S. Bezzateev, S. Andreev, Y. Koucheryavy, M. Gerla, Challenges of multi-factor authentication for securing advanced iot applications, IEEE Network 33 (2) (2019) 82–88.

[9] S. Gupta, A. Buriro, B. Crispo, Demystifying authentication concepts in smartphones: Ways and types to secure access, Mobile Information Systems 2018.

[10] J. Bonneau, C. Herley, P. C. Van Oorschot, F. Stajano, The quest to replace passwords: A framework for comparative evaluation of web authentication schemes, in: Proceedings of the Symposium on Security and Privacy, IEEE, 2012, pp. 553–567.

[11] A. M. Gamundani, A. Phillips, H. N. Muyingi, An overview of potential authentication threats and attacks on internet of things (iot): A focus on smart home applications, in: Proceedings of the International Conference on Internet of Things (iThings), IEEE, 2018, pp. 50–57.

[12] M. Antonakakis, Understanding the mirai botnet, in: Proceedings of the $26^{\text{th}}$ USENIX Security Symposium), 2017, pp. 1093–1110.

[13] C. Katsini, M. Belk, C. Fidas, N. Avouris, G. Samaras, Security and usability in knowledge-based user authentication: A review, in: Proceedings of the $20^{\text{th}}$ Pan-Hellenic Conference on Informatics, 2016, pp. 1–6.

[14] Namirial, Viewsonic to offer esignature solutions powered by namirial software, https://www.namirial.com/en/viewsonic-to-offer-esignature-solutions-powered-by-namirial-software/, online web resource (*Accessed on 30-06-2021*).

[15] R. Bhuyan, S. P. K. Kenny, S. Borah, D. Mishra, K. Das, Recent advancements in continuous authentication techniques for mobile-touchscreen-based devices, in: Proceedings of the Intelligent and Cloud Computing, Springer, 2021, pp. 263–273.

[16] BehavioSec, Behaviosec: Behavioral biometrics, https://www.behaviosec.com/wp-content/uploads/2018/09/Behaviosec-FAQ-Starting-point-09072018-ML-V2.pdf, online web resource (*Accessed on 30-06-2021*).

[17] BioCatch, End-to-end digital identity protection is now hassle-free, https://www.biocatch.com/behavioral-biometrics-cyber-security-software-tools, online web resource (*Accessed on 30-06-2021*).

[18] A. Tunnell, S. Powers, J. Zurasky, D. Tunnell, Biometric, behavioral-metric, knowledge-metric, and electronic-metric directed authentication and transaction method and system, uS Patent App. 15/202,515 (Jan. 11 2018).

[19] A. Ross, S. Banerjee, A. Chowdhury, Security in smart cities: A brief review of digital forensic schemes for biometric data, Pattern Recognition Letters 138 (2020) 346–354.

[20] S. Gupta, B. Crispo, A perspective study towards biometric-based rider authentication schemes for driverless taxis, in: Proceedings of the International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), IEEE, 2019, pp. 1–6.

[21] J. Wang, Y. Chen, S. Hao, X. Peng, L. Hu, Deep learning for sensor-based activity recognition: A survey, Pattern Recognition Letters 119 (2019) 3–11.

[22] NEC, A seamless curb-to-gate experience, https://www.nec.com/en/global/solutions/safety/aviation/experience/index.html, online web resource (*Accessed on 30-06-2021*).

[23] A. M. Research, Behavioral biometrics market outlook: 2025, https://www.alliedmarketresearch.com/behavioral-biometrics-market, online web resource (*Accessed on 30-06-2021*).

[24] R. V. Yampolskiy, V. Govindaraju, Behavioural biometrics: a survey and classification, International Journal of Biometrics 1 (1) (2008) 81–113.

[25] W. Meng, D. S. Wong, S. Furnell, J. Zhou, Surveying the development of biometric user authentication on mobile phones, IEEE Communications Surveys & Tutorials 17 (3) (2014) 1268–1293.

[26] A. Alzubaidi, J. Kalita, Authentication of smartphone users using behavioral biometrics, IEEE Communications Surveys & Tutorials 18 (3) (2016) 1998–2026.

[27] R. Oak, A literature survey on authentication using behavioural biometric techniques, Intelligent Computing and Information and Communication (2018) 173–181.

[28] L. M. Dang, K. Min, H. Wang, M. J. Piran, C. H. Lee, H. Moon, Sensor-based and vision-based human activity recognition: A comprehensive survey, Pattern Recognition 108 (2020) 107561.

[29] I. Stylios, S. Kokolakis, O. Thanou, S. Chatzis, Behavioral biometrics & continuous user authentication on mobile devices: A survey, Information Fusion.

[30] M. El-hajj, A. Fadlallah, M. Chamoun, A. Serhrouchni, A survey of internet of things (iot) authentication schemes, Sensors 19 (5) (2019) 1141.

[31] R. V. Yampolskiy, Behavioral, cognitive and virtual biometrics, in: Proceedings of the Computer Analysis of Human Behavior, Springer, 2011, pp. 347–385.

[32] Z. Hinbarji, Behavioural biometric identification based on human computer interaction, Ph.D. thesis, Dublin City University (2018).

[33] J. Spooren, D. Preuveneers, W. Joosen, Leveraging battery usage from mobile devices for active authentication, Mobile Information Systems 2017.

[34] D. Saravanan, Database security incursion recognition technique using neural network, in: Proceedings of the International Conference on Engineering and Technology Systems, Vol. 13, 2016, pp. 130–134.

[35] W. L. Al-Yaseen, Z. A. Othman, M. Z. A. Nazri, Real-time multi-agent system for an adaptive intrusion detection system, Pattern Recognition Letters 85 (2017) 56–64.

[36] N. Clarke, F. Li, A. Alruban, S. Furnell, Insider misuse identification using transparent biometrics, in: Proceedings of the 50$^{\text{th}}$ Hawaii International Conference on System Sciences, 2017, pp. 4031–4040.

[37] D. Brodić, A. Amelio, Human-computer interaction, in: Proceedings of The CAPTCHA: Perspectives and Challenges, Springer, 2020, pp. 7–14.

[38] L. Babula, K. Burda, User model for determining user's motor skills, Ph.D. thesis, Institute of Informatics, Information Systems and Software Engineering, FIIT STU Bratislava (2019).

[39] R.-D. Vatavu, Fundamentals of gesture production, recognition, and analysis, in: Proceedings of the CHI Conference Extended Abstracts on Human Factors in Computing Systems, 2017, pp. 1174–1177.

[40] J. G. Ulloa, Applied Biomechatronics Using Mathematical Models: Experiment design, data acquisition and signal processing, Academic press, 2018.

[41] S. Gupta, A. Buriro, B. Crispo, Smarthandle: A novel behavioral biometric-based authentication scheme for smart lock systems, in: Proceedings of the 3$^{\text{rd}}$ International Conference on Biometric Engineering and Applications, 2019, pp. 15–22.

[42] S. Gupta, A. Buriro, B. Crispo, Driverauth: Behavioral biometric-based driver authentication mechanism for on-demand ride and ridesharing infrastructure, ICT Express 5 (1) (2019) 16–20.

[43] B. Yang, Y. Zhang, Z. Liu, X. Jiang, M. Xu, Handwriting posture prediction based on unsupervised model, Pattern Recognition 100 (2020) 107093.

[44] M. Edwards, X. Xie, Footstep pressure signal analysis for human identification, in: Proceedings of the 7$^{\text{th}}$ International Conference on Biomedical Engineering and Informatics, IEEE, 2014, pp. 307–312.

[45] I. 2382-37:2017(en), Information technology — vocabulary — part 37: Biometrics, `https://www.iso.org/obp/ui/#iso:std:iso-iec:2382:-37:ed-2:v1:en`, online web resource (*(Accessed on 30-06-2021)*).

[46] ISO/IEC24741:2018(en), Information technology — biometrics — overview and application, `https://www.iso.org/obp/ui/#iso:std:iso-iec:tr:24741:ed-2:v1:en`, online web resource (*(Accessed on 30-06-2021)*).

[47] ISO/19795-1:2006(en), Biometric performance testing and reporting, `https://www.iso.org/obp/ui/fr/#iso:std:iso-iec:19795:-1:ed-1:v1:en`, online web resource (*(Accessed on 30-06-2021)*).

[48] M. Frank, R. Biedert, E. Ma, I. Martinovic, D. Song, Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication, IEEE transactions on information forensics and security 8 (1) (2012) 136–148.

[49] R. W. Soames, Anatomy and Human Movement E-Book: Structure and function, Elsevier Health Sciences, 2018.

[50] P. S. Teh, N. Zhang, A. B. J. Teoh, K. Chen, A survey on touch dynamics authentication in mobile devices, Computers & Security 59 (2016) 210–235.

[51] G. Kambourakis, D. Damopoulos, D. Papamartzivanos, E. Pavlidakis, Introducing touchstroke: keystroke-based authentication system for smartphones, Security and Communication Networks 9 (6) (2016) 542–554.

[52] N. Zheng, K. Bai, H. Huang, H. Wang, You are how you touch: User verification on smartphones via tapping behaviors, in: Proceedings of the $22^{\text{nd}}$ International Conference on Network Protocols, IEEE, 2014, pp. 221–232.

[53] P. S. Teh, N. Zhang, S.-Y. Tan, Q. Shi, W. H. Khoh, R. Nawaz, Strengthen user authentication on mobile devices by using user's touch dynamics pattern, Journal of Ambient Intelligence and Humanized Computing (2019) 1–21.

[54] A. Buriro, S. Gupta, A. Yautsiukhin, B. Crispo, Risk-driven behavioral biometric-based one-shot-cum-continuous user authentication scheme, Journal of Signal Processing Systems.

[55] A. K. Belman, V. V. Phoha, Discriminative power of typing features on desktops, tablets, and phones for user identification, Transactions on Privacy and Security (TOPS) 23 (1) (2020) 1–36.

[56] A. Jain, V. Kanhangad, Exploring orientation and accelerometer sensor data for personal authentication in smartphones using touchscreen gestures, Pattern recognition letters 68 (2015) 351–360.

[57] Y. Song, Z. Cai, Z.-L. Zhang, Multi-touch authentication using hand geometry and behavioral information, in: Proceedings of the IEEE Symposium on Security and Privacy (SP), IEEE, 2017, pp. 357–372.

[58] S. Gupta, A. Buriro, B. Crispo, Driverauth: A risk-based multi-modal biometric-based driver authentication scheme for ride-sharing platforms, Computers & Security 83 (2019) 122–139.

[59] E. Ellavarason, R. Guest, F. Deravi, Evaluation of stability of swipe gesture authentication across usage scenarios of mobile device, EURASIP Journal on Information Security 2020 (1) (2020) 1–14.

[60] W. Li, J. Tan, W. Meng, Y. Wang, A swipe-based unlocking mechanism with supervised learning on smartphones: Design and evaluation, Journal of Network and Computer Applications (2020) 102687.

[61] M. Antal, L. Z. Szabó, Biometric authentication based on touchscreen swipe patterns, Procedia Technology 22 (2016) 862–869.

[62] J. Fierrez-Aguilar, S. Krawczyk, J. Ortega-Garcia, A. K. Jain, Fusion of local and regional approaches for on-line signature verification, in: Proceedings of the International Workshop on Biometric Person Authentication, Springer, 2005, pp. 188–196.

[63] N. Li, J. Liu, Q. Li, X. Luo, J. Duan, Online signature verification based on biometric features, in: Proceedings of the $49^{\text{th}}$ Hawaii international conference on system sciences (HICSS), IEEE, 2016, pp. 5527–5534.

[64] R. Tolosana, R. Vera-Rodriguez, R. Guest, J. Fierrez, J. Ortega-Garcia, Exploiting complexity in pen-and touch-based signature biometrics, International Journal on Document Analysis and Recognition (IJDAR) (2020) 1–13.

[65] T. Yoshida, Y. Tanaka, S. Hangai, A study on signature/sign authentication with touching information on smart phone, in: Proceedings of the $9^{\text{th}}$ International Conference on Bioinformatics and Biomedical Technology, 2017, pp. 80–83.

[66] M. Anusuya, S. Katti, Speech recognition by machine: a review, International Journal of Computer Science and Information Security 3.

[67] L. Docio-Fernandez, C. García Mateo, Speech production, in: Proceedings of the Encyclopedia of Biometrics, Springer US, 2015, pp. 1493–1498.

[68] L. Zhang, S. Tan, J. Yang, Y. Chen, Voicelive: A phoneme localization based liveness detection for voice authentication on smartphones, in: Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, 2016, pp. 1080–1091.

[69] M. G. Christensen, Pitch estimation, in: Proceedings of the Introduction to Audio Processing, Springer, 2019, pp. 179–192.

[70] S. O. Sadjadi, J. H. Hansen, Mean hilbert envelope coefficients (mhec) for robust speaker and language identification, speech communication 72 (2015) 138–148.

[71] B. Kurian, V. Sreehari, L. Mary, Pncc for forensic automatic speaker recognition, in: Proceeding of the AIP Conference Proceedings, Vol. 2222, AIP Publishing LLC, 2020, p. 030004.

[72] C. Joder, B. Schuller, Exploring nonnegative matrix factorization for audio classification: Application to speaker recognition, in: Proceedings of the ITG Speech Communication Symposium, VDE, 2012, pp. 1–4.

[73] M. Baelde, C. Biernacki, R. Greff, Real-time monophonic and polyphonic audio classification from power spectra, Pattern Recognition 92 (2019) 82–92.

[74] M. A. Rao, P. K. Ghosh, Pitch prediction from mel-generalized cepstrum—a computationally efficient pitch modeling approach for speech synthesis, in: Proceedings of the $25^{\text{th}}$ European Signal Processing Conference (EUSIPCO), IEEE, 2017, pp. 1629–1633.

[75] T. S. Nguyen, K. Kilgour, M. Sperber, A. Waibel, Improved speaker adaptation by combining i-vector and fmllr with deep bottleneck networks, in: Proceedings of the International Conference on Speech and Computer, Springer, 2017, pp. 417–426.

[76] N. Maghsoodi, H. Sameti, H. Zeinali, T. Stafylakis, Speaker recognition with random digit strings using uncertainty normalized hmm-based i-vectors, Transactions on Audio, Speech, and Language Processing 27 (11) (2019) 1815–1825.

[77] S.-Y. Chang, N. Morgan, Robust cnn-based speech recognition with gabor filter kernels, in: Proceedings of the $15^{\text{th}}$ annual conference of the international speech communication association, 2014, pp. 1–5.

[78] Y. Rahulamathavan, K. R. Sutharsini, I. G. Ray, R. Lu, M. Rajarajan, Privacy-preserving ivector-based speaker verification, IEEE/ACM Transactions on Audio, Speech, and Language Processing 27 (3) (2018) 496–506.

[79] X. Yuan, G. Li, J. Han, D. Wang, Z. Tiankai, Speaker identification based on ivector and xvector, in: Proceedings of the Journal of Physics: Conference Series, Vol. 1827, IOP Publishing, 2021, p. 012133.

[80] M. S. Nixon, T. Tan, R. Chellappa, Human identification based on gait, Vol. 4, Springer Science & Business Media, 2010.

[81] R. Liao, S. Yu, W. An, Y. Huang, A model-based gait recognition method with body pose and human prior knowledge, Pattern Recognition 98 (2020) 107069.

[82] M. A. Laribi, S. Zeghloul, Human lower limb operation tracking via motion capture systems, in: Design and Operation of Human Locomotion Systems, Elsevier, 2020, pp. 83–107.

[83] A. Alamdari, V. N. Krovi, A review of computational musculoskeletal analysis of human lower extremities, in: Proceedings of the Human Modelling for Bio-Inspired Robotics, Elsevier, 2017, pp. 37–73.

[84] X. Wang, T. Yang, Y. Yu, R. Zhang, F. Guo, Footstep-identification system based on walking interval, IEEE Intelligent Systems 30 (2) (2015) 46–52.

[85] O. Costilla-Reyes, R. Vera-Rodriguez, P. Scully, K. B. Ozanyan, Analysis of spatio-temporal representations for robust footstep recognition with deep residual neural networks, Transactions on pattern analysis and machine intelligence 41 (2) (2018) 285–296.

[86] T. Van Nguyen, N. Sae-Bae, N. Memon, Draw-a-pin: Authentication using finger-drawn pin on touch devices, computers & security 66 (2017) 115–128.

[87] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, Biotouchpass: Handwritten passwords for touchscreen biometrics, IEEE Transactions on Mobile Computing.

[88] W. Li, W. Meng, S. Furnell, Exploring touch-based behavioral authentication on smartphone email applications in iot-enabled smart cities, Pattern Recognition Letters 144 (2021) 35–41.

[89] B. Lab, Biometrics and data pattern analytics, `http://atvs.ii.uam.es/atvs/e-BioDigit.html`, online web resource (*Accessed on 30-06-2021*).

[90] A. Pozo, J. Fierrez, M. Martinez-Diaz, J. Galbally, A. Morales, Exploring a statistical method for touchscreen swipe biometrics, in: Proceedings of the International Carnahan Conference on Security Technology (ICCST), IEEE, 2017, pp. 1–4.

[91] A. Garbuz, A. Epishkina, K. Kogos, Continuous authentication of smartphone users via swipes and taps analysis, in: Proceedings of the European Intelligence and Security Informatics Conference (EISIC), IEEE, 2019, pp. 48–53.

[92] S. Y. Ooi, A. B.-J. Teoh, Touch-stroke dynamics authentication using temporal regression forest, IEEE Signal Processing Letters 26 (7) (2019) 1001–1005.

[93] R. Kumar, V. V. Phoha, A. Serwadda, Continuous authentication of smartphone users by fusing typing, swiping, and phone movement patterns, in: Proceedings of the 8th international conference on biometrics theory, applications and systems (BTAS), IEEE, 2016, pp. 1–8.

[94] S. Gupta, A. Buriro, B. Crispo, A chimerical dataset combining physiological and behavioral biometric traits for reliable user authentication on smart devices and ecosystems, Data in brief 28 (2020) 104924.

[95] M. Frank, Touchalytics, `http://www.mariofrank.net/touchalytics/`, online web resource (*Accessed on 30-06-2021*).

[96] M. Gomez-Barrero, J. Galbally, J. Fierrez, J. Ortega-Garcia, R. Plamondon, Enhanced on-line signature verification based on skilled forgery detection using sigma-lognormal features, in: Proceedings of the international conference on biometrics (ICB), IEEE, 2015, pp. 501–506.

[97] Y. Ren, C. Wang, Y. Chen, M. C. Chuah, J. Yang, Signature verification using critical segments for securing mobile transactions, IEEE Transactions on Mobile Computing 19 (3) (2019) 724–739.

[98] M. M. Al-Jarrah, S. S. Al-Khafaji, S. Amin, X. Feng, Finger-drawn signature verification on touch devices using statistical anomaly detectors, in: Proceedings of the SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, IEEE, 2019, pp. 1700–1705.

[99] S. K. Behera, P. Kumar, D. P. Dogra, P. P. Roy, Fast signature spotting in continuous air writing, in: Proceeding of the 15th IAPR international conference on machine vision applications (MVA), IEEE, 2017, pp. 314–317.

[100] R. Ramachandra, S. Venkatesh, K. Raja, C. Busch, Handwritten signature and text based user verification using smartwatch, in: Proceedings of the 25th International Conference on Pattern Recognition (ICPR), IEEE, 2021, pp. 5099–5106.

[101] M. P. Centeno, A. van Moorsel, S. Castruccio, Smartphone continuous authentication using deep learning autoencoders, in: Proceedings of the 15th Annual Conference on Privacy, Security and Trust (PST), IEEE, 2017, pp. 147–1478.

[102] S. Amini, V. Noroozi, A. Pande, S. Gupte, P. S. Yu, C. Kanich, Deepauth: A framework for continuous user re-authentication in mobile apps, in: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, 2018, pp. 2027–2035.

[103] C. X. Lu, B. Du, X. Kan, H. Wen, A. Markham, N. Trigoni, Verinet: user verification on smartwatches via behavior biometrics, in: Proceedings of the First ACM Workshop on Mobile Crowdsensing Systems and Applications, 2017, pp. 68–73.

[104] A. Buriro, B. Crispo, M. Eskandri, S. Gupta, A. Mahboob, R. Van Acker, Snapauth: a gesture-based unobtrusive smartwatch user authentication scheme, in: Proceedings of the International Workshop on Emerging Technologies for Authorization and Authentication, Springer, 2018, pp. 30–37.

[105] B. Li, H. Sun, Y. Gao, V. V. Phoha, Z. Jin, Enhanced free-text keystroke continuous authentication based on dynamics of wrist motion, in: Proceeding of the IEEE Workshop on Information Forensics and Security (WIFS), IEEE, 2017, pp. 1–6.

[106] S. Mare, R. Rawassizadeh, R. Peterson, D. Kotz, Continuous smartphone authentication using wristbands, Workshop on Usable Security (USEC).

[107] J. Yang, Y. Li, M. Xie, Motionauth: Motion-based authentication for wrist worn smart devices, in: Proceedings of the International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops), IEEE, 2015, pp. 550–555.

[108] C. Bo, L. Zhang, X.-Y. Li, Q. Huang, Y. Wang, Silentsense: silent user identification via touch and movement behavioral biometrics, in: Proceedings of the 19th annual international conference on Mobile computing & networking, 2013, pp. 187–190.

[109] Z. Sitová, J. Šeděnka, Q. Yang, G. Peng, G. Zhou, P. Gasti, K. S. Balagani, Hmog: New behavioral biometric features for continuous authentication of smartphone users, IEEE Transactions on Information Forensics and Security 11 (5) (2015) 877–892.

[110] F. G. Barbosa, W. L. S. Silva, Support vector machines, mel-frequency cepstral coefficients and the discrete cosine transform applied on voice based biometric authentication, in: Proceedings of the SAI intelligent systems conference (IntelliSys), IEEE, 2015, pp. 1032–1039.

[111] L. Doddappagol, B. Geetha, User authentication using text-prompted technique, Asian Journal of Engineering and Technology Innovation 4 (7).

[112] J. Chauhan, Y. Hu, S. Seneviratne, A. Misra, A. Seneviratne, Y. Lee, Breathprint: Breathing acoustics-based user authentication, in: Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services, 2017, pp. 278–291.

[113] P. Musale, D. Baek, B. J. Choi, Lightweight gait based authentication technique for iot using subconscious level activities, in: Proceedings of the 4th World Forum on Internet of Things (WF-IoT), IEEE, 2018, pp. 564–567.

[114] D. Kastaniotis, I. Theodorakopoulos, C. Theoharatos, G. Economou, S. Fotopoulos, A framework for gait-based recognition using kinect, Pattern Recognition Letters 68 (2015) 327–335.

[115] D. Baek, P. Musale, J. Ryoo, Walk to show your identity: Gait-based seamless user authentication framework using deep neural network, in: Proceeding of the 5th ACM Workshop on Wearable Systems and Applications, 2019, pp. 53–58.

[116] P. Wasnik, K. Schafer, R. Ramachandra, C. Busch, K. Raja, Fusing biometric scores using subjective logic for gait recognition on smartphone, in: Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG), IEEE, 2017, pp. 1–5.

[117] I. Lamiche, G. Bin, Y. Jing, Z. Yu, A. Hadid, A continuous smartphone authentication method based on gait patterns and keystroke dynamics, Journal of Ambient Intelligence and Humanized Computing 10 (11) (2019) 4417–4430.

[118] W. Xu, Y. Shen, C. Luo, J. Li, W. Li, A. Y. Zomaya, Gait-watch: A gait-based context-aware authentication system for smart watch via sparse coding, Ad Hoc Networks (2020) 102218.

[119] R. Vera-Rodriguez, J. S. Mason, J. Fierrez, J. Ortega-Garcia, Comparative analysis and fusion of spatiotemporal information for footstep recognition, IEEE transactions on pattern analysis and machine intelligence 35 (4) (2013) 823–834.

[120] B. Zhou, M. S. Singh, S. Doda, M. Yildirim, J. Cheng, P. Lukowicz, The carpet knows: Identifying people in a smart environment from a single step, in: Proceedings of the International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), IEEE, 2017, pp. 527–532.

[121] J. E. Riwurohi, J. E. Istiyanto, K. Mustofa, A. E. Putra, People recognition through footstep sound using mfcc extraction method of artificial neural network backpropagation, International Journal of Computer Science and Network Security (IJCSNS) 18 (4) (2018) 28–35.

[122] S. Gupta, Next-generation user authentication schemes for iot applications, Ph.D. thesis, DISI, Univeristy of Trento, Italy (2020).

[123] A. Serwadda, V. V. Phoha, Z. Wang, R. Kumar, D. Shukla, Toward robotic robbery on the touch screen, ACM Transactions on Information and System Security (TISSEC) 18 (4) (2016) 1–25.

[124] Y. Li, M. Xie, Understanding secure and usable gestures for realtime motion based authentication, in: Proceedings of the Conference on Computer Communications Workshops (INFOCOM WKSHPS), IEEE, 2018, pp. 13–20.

[125] H. Feng, K. Fawaz, K. G. Shin, Continuous authentication for voice assistants, in: Proceedings of the $23^{rd}$ Annual International Conference on Mobile Computing and Networking, 2017, pp. 343–355.

[126] R. Rahmeni, A. B. Aicha, Y. B. Ayed, Speech spoofing countermeasures based on source voice analysis and machine learning techniques, Procedia Computer Science 159 (2019) 668–675.

[127] Y.-T. Chang, A two-layer authentication using voiceprint for voice assistants, Ph.D. thesis, University of Washington (2018).

[128] S. Garg, S. Bhilare, V. Kanhangad, Subband analysis for performance improvement of replay attack detection in speaker verification systems, in: Proceedings of the $5^{th}$ International Conference on Identity, Security, and Behavior Analysis (ISBA), IEEE, 2019, pp. 1–7.

[129] F. Tom, M. Jain, P. Dey, I. Kharagpur, End-to-end audio replay attack detection using deep convolutional networks with attention., in: Proceedings of the INTERSPEECH, 2018, pp. 681–685.

[130] M. Pal, D. Paul, G. Saha, Synthetic speech detection using fundamental frequency variation and spectral features, Computer Speech & Language 48 (2018) 31–50.

[131] J. Yang, R. K. Das, Long-term high frequency features for synthetic speech detection, Digital Signal Processing 97 (2020) 102622.

[132] Z. Wu, H. Li, On the study of replay and voice conversion attacks to text-dependent speaker verification, Multimedia Tools and Applications 75 (9) (2016) 5311–5327.

[133] M. Muaaz, R. Mayrhofer, Smartphone-based gait recognition: From authentication to imitation, IEEE Transactions on Mobile Computing 16 (11) (2017) 3209–3221.

[134] B. Shrestha, M. Mohamed, N. Saxena, Zemfa: Zero-effort multi-factor authentication based on multi-modal gait biometrics, in: Proceedings of the $17^{th}$ International Conference on Privacy, Security and Trust (PST), IEEE, 2019, pp. 1–10.

[135] L. Tran, T. Hoang, T. Nguyen, D. Choi, Improving gait cryptosystem security using gray code quantization and linear discriminant analysis, in: Proceedings of the International Conference on Information Security, Springer, 2017, pp. 214–229.

[136] Y. Yang, User behavior-based implicit authentication, Ph.D. thesis, University of Tennessee (2019).

[137] U. Burgbacher, K. Hinrichs, An implicit author verification system for text messages based on gesture typing biometrics, in: Proceedings of the SIGCHI conference on human factors in computing systems, 2014, pp. 2951–2954.

[138] A. Anitha, U. Gopalakrishnan, et al., A report on behavior-based implicit continuous biometric authentication for smart phone, in: Proceedings of Applied Computer Vision and Image Processing, Springer, 2020, pp. 169–184.

[139] J. Solano, L. Camacho, A. Correa, C. Deiro, J. Vargas, M. Ochoa, Combining behavioral biometrics and session context analytics to enhance risk-based static authentication in web applications, International Journal of Information Security (2020) 1–17.

[140] A. K. Jain, K. Nandakumar, A. Ross, 50 years of biometric research: Accomplishments, challenges, and opportunities, Pattern recognition letters 79 (2016) 80–105.

[141] C.-A. Toli, B. Preneel, Privacy-preserving biometric authentication model for e-finance applications., in: Proceedings of the ICISSP, 2018, pp. 353–360.

[142] EU, Principles of the gdpr, https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/principles-gdpr_en, online web resource (*Accessed on 30-06-2021*).

[143] C. Constitution, Ab-375 privacy: personal information: businesses, https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375, online web resource (*Accessed on 30-06-2021*).

[144] HHS, Summary of the hipaa privacy rule, https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html, online web resource (*Accessed on 30-06-2021*).

[145] ISO/IEC24745:2011(en), Biometric information protection, https://www.iso.org/obp/ui/#iso:std:iso-iec:24745:ed-1:v1:en, online web resource ((*Accessed on 30-06-2021*)).

[146] V. M. Patel, N. K. Ratha, R. Chellappa, Cancelable biometrics: A review, IEEE Signal Processing Magazine 32 (5) (2015) 54–65.

[147] E. Maiorana, P. Campisi, A. Neri, Bioconvolving: Cancelable templates for a multi-biometrics signature recognition system, in: Proceeding of the International Systems Conference, IEEE, 2011, pp. 495–500.

[148] J. Zhi, S. Y. Ooi, A. B. J. Teoh, Learning-based index-of-maximum hashing for touch-stroke template protection, in: Proceedings of the 12$^{th}$ International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), IEEE, 2019, pp. 1–6.

[149] K. Y. Chee, Design and analysis of voice template protection schemes based on winner-takes-all hashing, Ph.D. thesis, UTAR (2018).

[150] S. Billeb, C. Rathgeb, H. Reininger, K. Kasper, C. Busch, Biometric template protection for speaker recognition based on universal background models, IET Biometrics 4 (2) (2015) 116–126.

[151] L. A. Elrefaei, A. M. Al-Mohammadi, Machine vision gait-based biometric cryptosystem using a fuzzy commitment scheme, Journal of King Saud University-Computer and Information Sciences.

[152] E. A. Rúa, D. Preuveneers, W. Joosen, et al., Gait template protection using hmm-ubm, in: Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG), IEEE, 2018, pp. 1–8.

[153] ISO, Ergonomics of human-system interaction — part 11: Usability: Definitions and concepts, https://www.iso.org/standard/63500.html, online web resource (2018).

[154] J. D. Still, A. Cain, D. Schuster, Human-centered authentication guidelines, Information & Computer Security.

[155] S. Crispo, Bruno; Gupta, K. Halunen, Cybersec4europe: D3.7 usability requirements validation, https://cybersec4europe.eu/wp-content/uploads/2020/03/D3.7_Usability_requirements_validation_Submitted.pdf, online web resource (*Accessed on 30-06-2021*).
[156] S. Reis, A. Ferreira, P. M. V. Marques, R. Cruz-Correia, Usability study of a tool for patients' access control to their health data., in: Proceedings of the HEALTHINF, 2019, pp. 94–102.
[157] S. Dutta, Striking a balance between usability and cyber-security in iot devices, Ph.D. thesis, Massachusetts Institute of Technology (2017).
[158] J. M. Ferreira, S. T. Acuña, O. Dieste, S. Vegas, A. Santos, F. Rodríguez, N. Juristo, Impact of usability mechanisms: An experiment on efficiency, effectiveness and user satisfaction, Information and Software Technology 117 (2020) 106195.
[159] C. Wijayarathna, N. A. G. Arachchilage, Using cognitive dimensions to evaluate the usability of security apis: An empirical investigation, Information and Software Technology 115 (2019) 5–19.
[160] K. Halunen, J. Häikiö, V. Vallivaara, Evaluation of user authentication methods in the gadget-free world, Pervasive and Mobile Computing 40 (2017) 220–241.
[161] Y. X. M. Tan, A. Iacovazzi, I. Homoliak, Y. Elovici, A. Binder, Adversarial attacks on remote user authentication using behavioural mouse dynamics, in: Proceedings of the International Joint Conference on Neural Networks (IJCNN), IEEE, 2019, pp. 1–10.
[162] A. Krausova, Online behavior recognition: Can we consider it biometric data under gdpr, Masaryk UJL & Tech. 12 (2018) 161.
[163] G. Schumacher, Behavioural biometrics: Emerging trends and ethical risks, in: Proceedings of the Second Generation Biometrics: The Ethical, Legal and Social Context, Springer, 2012, pp. 215–227.
[164] F. Cherifi, B. Hemery, R. Giot, M. Pasquet, C. Rosenberger, Performance evaluation of behavioral biometric systems, in: Proceedings of Behavioral Biometrics for Human Identification: Intelligent Applications, IGI Global, 2010, pp. 57–74.
[165] IBIA, Behavioral biometrics, https://www.ibia.org/biometrics-and-identity/biometric-technologies/behavioral-biometrics, online web resource (*Accessed on 30-06-2021*).
[166] Y. Liang, S. Samtani, B. Guo, Z. Yu, Behavioral biometrics for continuous authentication in the internet of things era: An artificial intelligence perspective, IEEE Internet of Things Journal.
[167] R. D. Christ, R. L. Wernli, Chapter 12 - sensor theory, in: Proceedings of the ROV Manual (Second Edition), Butterworth-Heinemann, 2014, pp. 297 – 326.
[168] ST, Mems and sensors, https://www.st.com/en/mems-and-sensors.html, online web resource (*Accessed on 30-06-2021*).
[169] R. Semiconductor, Introducing the rokix sensor node, https://www.rohm.com/news-detail?news-title=roki-sensor-node&defaultGroupId=false, online web resource (*Accessed on 30-06-2021*).
[170] Topwaydisplay, Lcd touch screen comparison, https://www.topwaydisplay.com/blogs/lcd-touch-screen-comparison, online web resource (*Accessed on 30-06-2021*).
[171] S. Gao, J. Chen, Y. Dai, R. Wang, S. Kang, L. Xu, Piezoelectric based insole force sensing for gait analysis in the internet of health things, IEEE Consumer Electronics Magazine.
[172] A. Devices, Analog devices 3d time of flight (3d tof), https://www.analog.com/en/applications/technology/3d-time-of-flight.html#, online web resource (*Accessed on 30-06-2021*).
[173] Broadcom, Time-of-flight 3d sensors, https://www.broadcom.com/products/optical-sensors/time-of-flight-3d-sensors, online web resource (*Accessed on 30-06-2021*).

[174] Apple, Core motion, https://developer.apple.com/documentation/coremotion, online web resource (*Accessed on 30-06-2021*).

[175] Microsoft, Introduction to the sensor and location platform in windows, https://docs.microsoft.com/en-us/windows-hardware/drivers/sensors/, online web resource (*Accessed on 30-06-2021*).

[176] Android, Sensors overview, https://developer.android.com/guide/topics/sensors/sensors_overview, online web resource (*Accessed on 30-06-2021*).

[177] Intel, Intel ai hardware, https://www.intel.com/content/www/us/en/artificial-intelligence/hardware.html, online web resource (*Accessed on 30-06-2021*).

[178] M. Learning, Intel, https://www.intel.com/content/www/us/en/products/docs/storage/programmable/applications/machine-learning.html, online web resource (*Accessed on 30-06-2021*).