Behavioral Biometrics for Continuous Authentication in the Internet-of-Things Era: An Artificial Intelligence Perspective

Yunji Liang, Sagar Samtani, Bin Guo, and Zhiwen Yu, *Senior Member, IEEE*

***Abstract*—In the Internet-of-Things (IoT) era, user authentication is essential to ensure the security of connected devices and the customization of passive services. However, conventional knowledge-based and physiological biometric-based authentication systems (e.g., password, face recognition, and fingerprints) are susceptible to shoulder surfing attacks, smudge attacks, and heat attacks. The powerful sensing capabilities of IoT devices, including smartphones, wearables, robots, and autonomous vehicles enable continuous authentication (CA) based on behavioral biometrics. The artificial intelligence (AI) approaches hold significant promise in sifting through large volumes of heterogeneous biometrics data to offer unprecedented user authentication and user identification capabilities. In this survey article, we outline the nature of CA in IoT applications, highlight the key behavioral signals, and summarize the extant solutions from an AI perspective. Based on our systematic and comprehensive analysis, we discuss the challenges and promising future directions to guide the next generation of AI-based CA research.**

***Index Terms*—Artificial intelligence (AI), behavioral biometric, body area networks, constrained devices, continuous authentication (CA), cyber–physical systems data mining, Internet of Things (IoT).**

I. Introduction

With the flourishing of the Internet of Things (IoT), our daily life is being transformed by ambient intelligence [1] along with massively connected IoT devices ranging from smartphones and wearables to robots, autonomous vehicles, and drones [2], [3]. The broad penetration of IoT devices in the consumer market makes user authentication critically important to secure users have the appropriate right to access IoT devices [2] and to avoid the devastating damages caused by one attack occurring in the local vulnerable spots [4]. Apart from the security concerns, user authentication is beneficial for passive and customized services when user switching occurs. For example, for one autonomous car shared among family members, the driving habits among family members differ significantly. To assist the drivers, different assistance strategies can be applied based on user identities [3]. Thus, user authentication can protect crucial information against potential attacks and offer customized services for improved user experience.

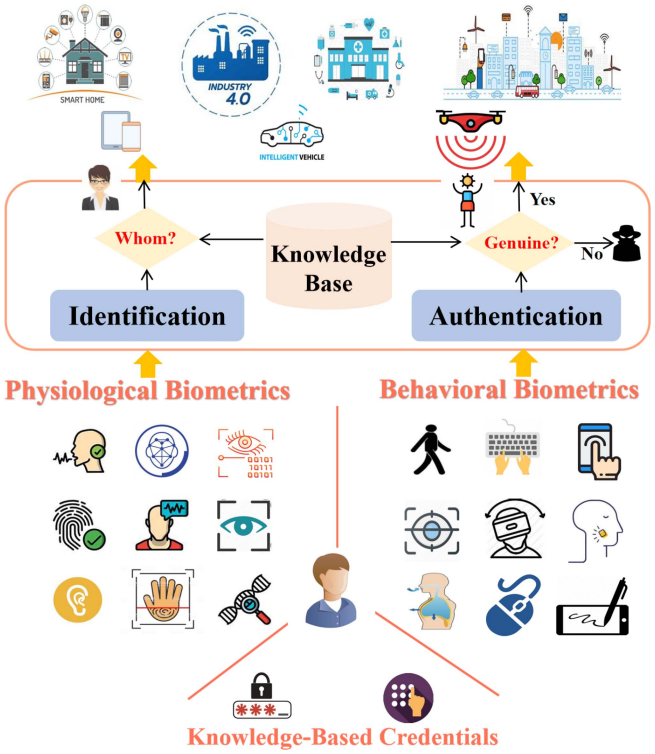

Fig. 1. Overview of credentials for user authentication and identification and their applications.

Due to the importance of user authentication, researchers and industries are increasingly studying the development of sophisticated methods to verify and recognize user identities. As shown in Fig. 1, authentication systems can be divided into three categories: 1) knowledge-based; 2) physiological biometric-based; and 3) behavioral biometric-based solutions [2], [5]. Knowledge-based authentication explicitly requests the user to enter credentials, such as password,

Manuscript received March 2, 2020; revised May 26, 2020; accepted June 16, 2020. Date of publication June 22, 2020; date of current version September 15, 2020. This work was supported in part by the National Key Research and Development Program of China under Grant 2019YFB2102200; in part by the Ministry of Health of China under Grant 2017ZX10303401-002 and Grant 2017YFC1200302; in part by the Natural Science Foundation of China under Grant 61902320, Grant 71472175, Grant 71602184, and Grant 71621002; in part by the National Science Foundation under Grant CNS-1850362 and Grant OAC-1917117; and in part by the Fundamental Research Funds for the Central Universities under Grant 31020180QD140. *(Corresponding author: Yunji Liang.)*

Yunji Liang, Bin Guo, and Zhiwen Yu are with the School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: liangyunji@nwpu.edu.cn).

Sagar Samtani is with the Operations and Decision Technologies Department, Kelley School of Business, Indiana University, Tampa, FL 33602 USA.

Digital Object Identifier 10.1109/JIOT.2020.3004077

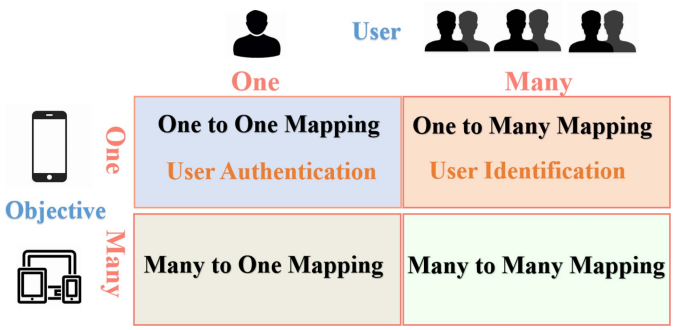

Fig. 2. Mapping relationship between devices and users.

personal identification number (PIN), and graphical PIN to confirm the identity of an individual. Physiological biometric-based authentication uses biological traits (e.g., fingerprint, iris, and facial images) and employs the machine learning methods to discriminate user identities. Behavioral biometrics, including walking gait, keystroke, and touchscreen dynamics are used for user authentication as well. Authentication systems can be classified into two subcategories: 1) *user authentication*, to detect whether the user is one unauthorized visitor or genuine user and 2) *user identification*, to recognize whom the current user is.

The essence of authentication systems is to build the mapping relationship between users and objectives. According to the object-user mapping relationship, authentication systems can be categorized as Fig. 2. Among them, one-to-one mapping aims to verify whether the user is a genuine user or imposter for one privately owned device (such as mobile phones and laptops) or one mobile application. One-to-many mapping provides the appropriate access control among multiple users for one object shared within a group of persons. In IoT systems, numerous smart devices are connected to provide pervasive services for one user (such as smart home and vehicle-to-vehicle systems [6]). In the dynamic environment, participants need to finish one session across shared IoT devices where complex and robust authentication schemes are needed [7]. The many-to-one mapping and many-to-many mapping fit well for the user authentication in the complex dynamic environment.

Although numerous user authentication and identification methods are proposed, prior methods have several key drawbacks as it pertains to their fit with IoT applications.

Vulnerability: Prior systems are prone to a diverse range of attacks. For knowledge-based authentication, imposters can capture inputs by shoulder surfing and recording attacks [8]–[10], thermal attack [11], [12], and smudge attacks [13], [14]. For facial recognition, an adversary could conquer the facial detection through legitimate users' facial photos. The fingerprint can be conquered by smudge attack [13]–[15] and forged by deep learning methods [16]. The automated speaker verification based on the personal characteristics of voices is subject to replay attacks [17], [18].

Discreteness: In general, user identification and authentication are executed once at the beginning of a session. If the authentication information is stolen or compromised, imposters can fully control the hacked accounts or IoT devices, resulting in devastating damages consequentially. In addition, one-time user authentication is insecure in some scenarios. For example, in the ride-sharing platforms, registered drivers may subcontract ride assignments or share their registration to an unauthorized person, which could be dangerous for the riders [3]. Thus, the one-time authentication method is insecure and cannot provide seamless protection.

Obtrusiveness: Existing solutions require explicit inputs or actions, which are obtrusive for users by requiring extra user attention. They also cause a distraction from the undergoing tasks [19], [20]. For example, iris and facial recognition require users to stare at the camera in specific angles, which is unnatural and uncomfortable for users.

In recent years, the rapid proliferation of IoT devices, such as smartphones, wearable devices, and facility cameras has made it possible to seamlessly sense and track user behaviors. The analysis and mining of behavior fingerprints offer new opportunities for continuous authentication (CA) [21], [22]. In this article, we provide a systematic overview of the CA based on behavioral biometrics from the perspective of artificial intelligence (AI). Our contributions in this article are as follows.

1) We provide a systematic overview of the key components and differentiators between user authentication and identification.
2) We summarize the key elements of behavioral biometrics.
3) We provide a summary of the emerging types of sensing technologies being integrated into emerging IoT technologies, with a specific focus on how the data they generate and common representations of these data.
4) We present a general framework on how future researchers can develop innovative AI-based approaches for continuous user authentication and identification.
5) We summarize emerging directions for future AI-based research in the aforementioned areas.

The remainder of this article is organized as follows. In Section II, we characterize the nature of behavioral biometrics. In Section III, we propose a general framework for continuous user authentication from sensing and computing perspectives. Sections IV and V provide one systematic survey about data sensing and inference methods, respectively. Finally, Section VI presents the open issues and challenges in CA based on behavioral biometrics, and Section VII concludes this article.

II. Characterizing Behavioral Biometrics

Behavioral biometrics refer to the unique behavioral traits that can be used for human authentication. Unlike the knowledge-based credentials and physiological biometrics shown in Fig. 1, behavioral biometrics identify people by how a user conducts the specified activity rather than by static information or physical characteristics. User authentication based on behavioral biometrics is characterized as *secure*, *continuous*, *transparent*, and *cost effective*.

Secure: In contrast to knowledge-based credentials and physiological biometrics, behavioral biometrics provide a

dynamic modality that is completely passive and works in the background, making it impossible to copy or steal. Behavioral biometric data are extracted when users are performing one specified activity. Unlike the static authentication information, the nature of behavioral biometric data ensures that they cannot be forgotten, exchanged, and stolen. Moreover, the dynamics of activities make it very difficult to forge behavioral biometrics. Authentication systems based on knowledge-based credentials and physiological biometrics are vulnerable to a variety of cyberattacks, including shoulder surfing attacks, smudge attacks, replay attacks, thermal attacks, and adversary attacks [8]–[10], [13], [14]. Systems based on behavioral biometrics are secure and robust to the aforementioned cyberattacks.

Continuous: In the IoT era, user authentication is one crucial task to secure the connected devices and it should not be a one-off event but rather a constant process. Unlike the one-time authentication that is enforced at the beginning of a session or login, continuous user authentication is an essential requirement to verify that users are who they claim to be on an ongoing basis. In order to achieve this goal, behavioral biometrics continuously profile a user's behavior based upon the natural interactions without having to constantly interrupt users. The continuity of behavior makes it a natural way for CA with no distraction for users.

Unobtrusive: Unobtrusive sensing aims to monitor physical activities and behaviors continuously via sensors embedded in the ambient environment or wearable sensors [23], and maximize the user experience to avoid disturbing users from the undergoing tasks [24]. Behavioral data can be sampled when users are interacting with IoT devices or ambient environments with no explicit input. Moreover, user authentication can be performed in a transparent and unobtrusive way with no distraction for users [25]. Previous attempts to continuously authenticate may have been too disruptive (e.g., prompts mid session), but now by using unobtrusive sensing techniques users can be continuously authenticated without interruption. This feature is beneficial for the enhancement of user experience and provides more secure protection for IoT devices.

Cost Effective: Physiological biometrics usually rely on customized hardware for information acquisition [26], [27]. This is often expensive in terms of costs and impedes the widespread adoption of physiological biometrics for user authentication. In contrast, behavioral biometrics can be observed and sampled with embedded sensors in IoT devices (e.g., microphone, touchscreen, accelerometer in smartphone, and wearable devices) or public facilities (e.g., WiFi access point and surveillance camera). The widespread availability of IoT devices makes it possible to sense behaviors without extra hardware, which improves the acceptance of behavioral biometrics with low cost and ease of use.

III. Overview of Continuous Authentication

To guide the readers to understand the core concepts in this survey article, we provide an overview of CA systems to illustrate what components should be included in the one behavior-based CA system. Specifically, we present an abstract framework of CA and identification based on behavioral signals to highlight the primary components for user authentication. As shown in Fig. 3, it consists of four layers: 1) *behavioral signals*; 2) *unobtrusive sensing*; 3) *continuous computing*; and 4) *applications*.

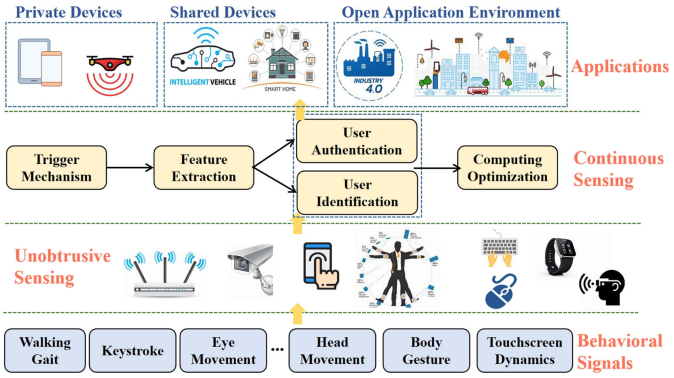

Fig. 3. Abstract framework of CA and identification based on behavioral signals.

Behavioral signals are the collection of distinctive behavioral patterns or traits that can be used by one decision-making system to decide an individual's identity. A large number of pilot studies show that dynamics of a keystroke, walking gaits, eye movements, and touchscreen dynamics are suitable for CA.

Unobtrusive sensing summarizes the available sensors and feasible sensing strategies to capture the behavioral signals. The sensing modalities for behavioral signals are diverse. In the IoT era, the unprecedented sensing capability brings opportunities to sense the behavior in different granularities with diverse sensors. For example, walking gaits can be captured by facility cameras, accelerometers built-in wearables, and WiFi signals that bounceoff the walking individuals. IoT devices, such as smartphones and wearables are infused into our daily life and can provide transparent, unobtrusive, and continuous behavior sensing without additional attentions and actions required.

Continuous computing highlights the workflow for CA. The goals of CA based on behavioral biometrics are to detect whether the user has the right to access the IoT device or not (*authentication*) and to recognize who the current user is (*identification*). Accordingly, CA based on behavioral biometrics can be divided into two categories: 1) anomaly detection and 2) classification. Anomaly detection methods can determine the abnormal patterns from the regular ones. For user identification, one predictive model is trained to maximize the interclass differences (i.e., legitimate users versus outliers).

Applications are typical scenarios where CA based on behavioral biometrics is applied. According to the property of devices, user authentication can be applied in three categories of IoT devices or scenarios, including private devices, shared devices, and open application environments. Apart from IoT device security issues for access control, behavioral biometrics can be applied for customized services in smart space. For example, for one autonomous car shared among family members, insurance companies can design tailored insurance

policies according to individual's driving patterns, where the continuous user identification based on driving patterns is the premise. For smart spaces, understanding the presence of users in the buildings is significantly important for providing more responsive and customized services [28].

IV. Sensing of Behavioral Biometrics

Numerous behavioral traits have been explored for CA. In this section, we analyze the commonly used behavioral traits for use authentication, and conduct comprehensive comparison from different dimensions, including *vulnerability*, *discretenss*, *obtrusiveness*, and *privacy*.

A. Keystroke Dynamics

Keystroke dynamics characterize the typing rhythm, such as keystroke length, the distance between consecutive strokes, the pressure exerted on each key when the individual types characters, and others. To date, keystroke-powered authentication has been broadly explored for devices equipped with physical keyboards. With the emergence of touchscreens, when users enter characters via touchscreens, subtle changes of built-in sensors, including accelerometer and gyroscope occur. Jointly combined with the status of built-in sensors, the keystroke timing, touch-typing, and keystroke pressure are distinctive features for user identification [29], [30]. Similarly, mouse usage dynamics have also been shown to serve as potential authentication cues [31], [32].

The advantages of analyzing keystroke dynamics include the unobtrusive data collection and continuous monitoring of typing behaviors when users interact with devices simultaneously. However, the keystroke dynamics vary in different scenarios, such as walking, holding at hand, and putting on table [33]. As a result, keystroke-based user authentication is scenario dependent, which requires the understanding of user scenarios and build the appropriate algorithms accordingly.

B. Touchscreen Dynamics

With the prevalence of touchscreen in IoT devices, sophisticated interaction patterns, including pressure intensity and sliding dynamics when users interact with touchscreen enable the detection of user identification in an unobtrusive way [34]. One type of study combines touch patterns with the conventional authentication method, such as PIN codes or shaped-based drawing when individuals are running the log-in session. Even though more patterns are extracted to protect the devices against potential attacks, the authentication method is still static. For CA, Sitová *et al.* [22] analyzed the micromovement and orientation dynamics resulting from how a user grasps, holds, and taps on the smartphone and leveraged the context and touchscreen dynamics for user authentication.

The authentication based on touch operations provides one natural way to collect user interaction data. Moreover, it makes continuous user identification possible with no extra sensors and low computational load. However, touch operations vary among applications. Therefore, systematically studying application-dependent touch patterns can help protect against unauthorized access of crucial mobile applications. This is especially true when individuals are likely to possess more than one mobile device. When they interact with different devices, whether the touch dynamics are identical and can be transferred among different devices are still open questions.

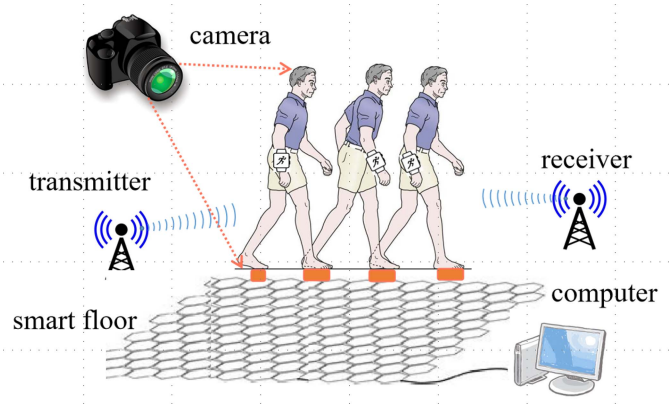

Fig. 4. User authentication and identification based on walking gaits via different sensors, including cameras, wearables, smart floor, and device-free sensing.

C. Eye Movement

Driven by the internal interaction relationship between muscles and brain neural, eye movements, including gaze and blinking are significantly different for individuals and are difficult to be mimicked and duplicated. Authentication based on eye movements can be divided into two categories in terms of data signals. Bioelectrical signals caused by eye movements and blinks are studied and found that the accompanying electrooculogram signals extracted from eye blinking were unique and rational as the biometrics for identification recognition tasks [35]. The dynamics of eye movements, including pupillary response to stimuli, pupil size, velocity, acceleration, and spatial/geometric features are recorded and analyzed from the video. Several studies that demonstrated those patterns were intrinsic and could be applied for user identification [36]–[38].

However, eye-movement-based solutions often use expensive and specialized monitor-mounted gaze trackers. Such explicit authentication methods may cause vigilance of the imposter and cover the camera of the tracker. These authentication methods also have high energy consumption for continuous video recording and real-time video analysis. Furthermore, eye-movement-based user identification can be obtrusive to an individual's privacy.

D. Walking Gait

Identifying and authenticating based on walking gaits are an emerging biometric technology that recognizes users' identities by analyzing walking patterns [25]. Based on the strategies of data acquisition, the sensing strategies of gait signals can be grouped as: facility cameras, floor sensors, and wearables. Fig. 4 illustrates the interplay and relationships among the three components.

Vision-based solutions record an individual's gait patterns when walking via facility cameras. Then, background segmentation techniques are used to extract features from recorded

images to verify user identities [39], [40]. However, the vision-based solutions are subject to environments, including illumination and camera angle [39], [41]. Furthermore, the high computation consumption and privacy concerns make vision-based solutions infeasible for CA.

For floor-sensor based solutions, dense press sensors are deployed under the floor to track the pressure dynamics or acoustic patterns when walking on the floor [42]. Its advantages include high resolution in terms of performance and unobtrusiveness for user interaction [43]. However, floor-sensor-based solutions are ideal for CA for two reasons. First, they often have sophisticated system design and high costs. Second, they only work in an enclosed environment with limited users and do not work in the open space with low scalability. Taken together, these limitations often result in floor-based systems being difficult to deploy.

Wearable sensor-based solutions rely on sensors attached to different spots of the body (such as waist, hip, and pocket) to capture the accompanying signals when walking, therefore, enabling continuous verification of user identity [44]–[48]. However, primary studies are conducted in the laboratories with cumbersome prototype systems and expensive customized devices. Recently, more studies focus on user verification based on off-the-shelf devices (e.g., mobile phones) [49]–[51]. The main advantage of using a wearable accelerometer sensor for gait recognition is that it provides unobtrusive verification without requiring user explicit actions. Especially, the accelerometer sensor has characters of small volume, low cost, and can be easily integrated into the hardware of wearable devices.

Lately, walking gait recognition based on WiFi, millimeter wave, and radio-frequency identification (RFID) is gaining attention [52]. These studies assume that the unique walking gaits and body shapes entail distinctive disturbances in signals that can be used for user verification [53]. Wang *et al.* [54] provided one comprehensive survey about the device-free sensing based on WiFi signals. Several user verification systems based on WiFi signals are available, including WiFiID [55], WiWho [56], WiFiU [57], and FreeSense [58]. Among them, FreeSense [58] is an unobtrusive system for indoor human identification based on the disturbed WiFi channel-state information (CSI) signals when individuals walk through the line-of-sight (LOS) path between the source and the receiver of WiFi signals. FreeSense captures the disturbing waveforms when the user is walking across the LOS path, and applies the discrete wavelet transform and principal component analysis to extract shape features. The performance of FreeSense declines from 94.5% to 75.5% while the number of participants increases from two to nine. Luo *et al.* [52] used RFID for gait recognition by monitoring the interruptions to RFID signals when one target user is blocking the signals between transmitters and receivers.

However, CA based on walking gaits has several challenges. First, prior studies were mainly conducted in a controlled environment. Robustness should be further evaluated in the physical world. Second, a person's walking gaits can be altered by many factors, i.e., drunkenness, aging, carrying a load, and shoe type [59]–[61]. The model trained with the data set collected in one situation may introduce bias when it is applied to other situations [62]. Third, prior studies mainly focus on single-person gait detection. However, vision-based, floor-based, and WiFi-based solutions achieve suboptimal performances for multiperson scenarios.

E. Body Gesture

With the popularity of wearable sensors, body gestures especially hand gestures have been widely studied for user authentication. The majority of user authentication based on body gestures attempt to verify or recognize the user identity based on a specified gesture performed. Among them, Li *et al.* [66] found that a person's head movement patterns are unique when stimulated by music beats, and implemented the Headbanger, an authentication system that can authenticate uses by sensing head movements when listening to music beats based on the built-in accelerator in Google glasses. In addition, hand gesture and in-air handwriting are studied for user authentication as well. Matsuo *et al.* [65] designed one authentication system based on the acceleration signals during the arm sweep action. Lu *et al.* [68] proposed a multifactor user authentication framework using both the motion signal of a piece of in-air handwriting and the geometry of the hand skeleton captured by a depth camera. However, user authentication based on body gestures is obtrusive as the users are required to perform certain movements or actions.

F. Others

Chewing renders the changes of muscle tension with accompanying chewing sounds. Zou *et al.* [69] proposed a human authentication mechanism that utilized the sounds generated by dental occlusion (i.e., tooth click) to unlock the mobile devices. The prototype system, BiLock, relies on the microphone in mobile devices to record the sounds of dental occlusion and verifies whether the current user is legitimate or not. Although BiLock is easy to use and requires no extra sensing unit, its performance is sensitive to the scenario. Furthermore, BiLock requires users to put the mobile device 5–15 cm away from lips, which is obtrusive for users and does not support transparent sensing. Similarly, Bodybeat leverages the nonspeech body sounds caused by food intake, breath, laughter, and cough for user identification [70].

In addition, breathing is used for user authentication as well by characterizing the subtle vibration caused by respiratory. For example, BreathPrint employs deep-learning models to effectively express the acoustic features caused by breathing for user authentication on resource-constrained devices [71]. To provide unobtrusive CA, Liu *et al.* [72] proposed a continuous user verification system based on unique human respiratory-biometric characteristics extracted from the off-the-shelf WiFi signals.

G. Summary of Behavioral Biometrics

As shown in Table I, each behavioral biometric has its pros and cons, and no single biometric is expected to effectively meet all the needs of any scenarios and applications [73]. The advantages of most user authentication systems based on

TABLE I
SUMMARY OF BEHAVIORAL SIGNALS FOR USER AUTHENTICATION

Category	Signal	Description	Vulnerability	Discreteness	Obtrusiveness	Privacy
Keystroke Dynamics	typing [29]–[32]	typing characters via keyboard or clicking via touchscreen when entering information	√	×	×	√
Touch Operation	screen touch [22], [34]	finger movements and screen pressure on touchscreen	√	×	×	√
Eye Movement	biosignal [35]	bioelectronic signals caused by eye movements	×	×	√	√
	temporal [36]–[38]	dynamics of eye movement	√	×	√	√
Walking Gait	camera [39], [40]	capture the walking styles from video	√	×	×	√
	floor sensor [42], [43]	features caused by users walking on the pressure sensors embedded in floor	×	×	×	×
	wearable [44]–[48]	motions patterns tracked by wearable sensors	×	×	√	√
	WiFi [54], [55]	Perturbations caused by unique walk styles when walking through the WiFi spectrum field	×	×	×	√
	millimeter wave [63], [64]	interrupted radar signals between transmit antennas and receivers	×	×	×	√
	RFID [52]	Perturbations caused by unique walk styles when walking through the RFID spectrum field	×	×	×	√
Body Gesture	arm swing [65]	motion signals generated when a fixed behavior is performed	×	×	√	√
	head movement [66]	head motions measured by head-worn devices	×	×	×	√
	mouse movement [67]	Doppler profiles of acoustic signals caused by users' speaking mouths	×	×	×	√
	handwriting [68]	motion signals generated by in-air-handwriting	×	√	√	√
Sounds of Behavior	dental occlusion [69]	unique acoustic patterns of teeth click	√	√	√	√
	body sounds [70]	acoustic patterns caused by food intake, breath, and laugh	×	×	√	√
	breathing [71], [72]	acoustic features caused by breathing	×	×	×	√

behavioral biometrics include nonobtrusiveness with no extra user attention required and nonvulnerability against cyberattacks. However, the existing authentication systems do not take advantage of the nature of behavioral traits to support the CA and do not consider the issue of privacy protection.

V. AI-BASED SOLUTIONS FOR USER AUTHENTICATION

In this section, we summarize the AI-based methods that are employed to recognize user identities based on behavioral fingerprints. In CA, intelligent algorithms, including machine learning and deep learning are capable of determining the access control of IoT devices by checking user identities. As shown in Fig. 5, the pipeline of AI-based methods for CA consists of the following major components: *data preprocessing*, *feature extraction*, and *classification algorithms*. We further describe each component in the following sections.

A. Data Preprocessing

Data preprocessing is a critical procedure to distill high-quality data out of the raw data that are generally incomplete, noisy, inconsistent, and redundant. As the inputs of continuous human identification based on behavioral patterns are sequential data, data filtering and data segmentation are necessary to reduce the noisy data and align the inputs.

For data filtering, numerous filters are applied to the sequential data to remove data that can be repetitive, irrelevant, or even sensitive [55], [69]. For example, a Butterworth filter is applied to WiFi CSI data to remove the high-frequency noisy

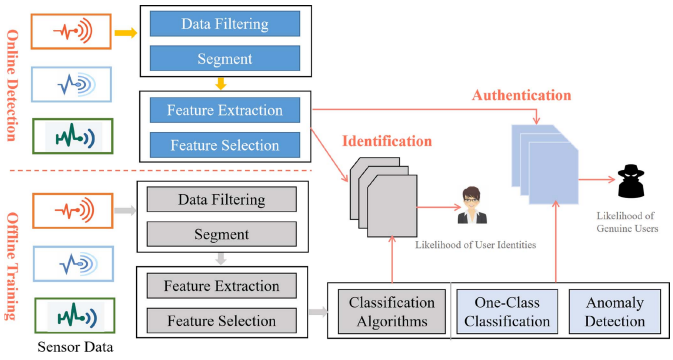

Fig. 5. General workflow of machine learning-based user identification and authentication.

data [55], [56]. Similarly, BiLock uses a six-order Butterworth filter to remove the out-of-band interference of dental clicks, and employs wavelet denoising to improve the signal to noise radio [69].

For the CA, data segmentation seeks effective regions from sequential data. The rule-based solutions, including fixed threshold or fixed-size windows are used over sequential data for segmentation. In WiFi-ID [55], two frequency bands are used to separate WiFi signals impacted by walking gaits from the wavelet domain. The fixed-size window is applied [74], [75] to segment the dynamic swipe behavior and video records of walking gaits into fixed-length slices, respectively. However, rule-based solutions are sensitive to inputs and tightly rely on prior knowledge. Dynamic segmentation strategies are

introduced to split the stream data dynamically. For example, BehaveSense [76] recognizes four touchscreen operations to separate the effective samples from stream data and utilizes the sequential patterns of touchscreen operations for user identification. In addition, dynamic time warping (DTW) is widely used to find out the cycle of behaviors [56], [58], [66].

B. Feature Extraction

Feature engineering is to extract features of value that can represent users' behavior comprehensively from the training data set. The extracted features depend on sensors and applications. Statistical features refer to the measurements of interpreting both quantitative and qualitative data with standard statistics, such as the root mean square, mean, standard deviation, and variance. Statistical features characterize the overall patterns of the given samples from macroperspectives. Due to the dynamics of user behaviors over time, frequency-domain patterns of behaviors are of significance for describing the dynamics of signals. To obtain frequency-domain patterns, sampled data are transformed by spectrum analysis to learn the frequency-domain features. As shown in Table II, DeepAuth characterizes the frequency-domain representation of motion sensors [74]. GlassGuard [77] and BiLock [69] employ the mel-frequency cepstral coefficients (MFCCs) to characterize the dynamics of vocal signals.

Generally, hand-crafted feature engineering heavily relies on the knowledge of domain experts and is time consuming to construct one complete feature set. As a result, it is the bottleneck of classification-oriented tasks. In addition, as not all hand-crafted features are preeminently contributive to the verification of user identity, feature selection is optional to rebuild one subset of attributes with least data loss. To mitigate the behavioral variability of mouse dynamics, Cai *et al.* [81] proposed a unified framework of employing dimensionality reduction methods to extract predominant characteristics from the original feature space for enhanced performance and found that variability reduction in feature engineering could enhance the authentication mechanisms. On the other hand, due to these drawbacks of feature engineering, many researchers in this field have turned to deep learning-based methods. Unlike conventional machine learning that relies on feature engineering, deep learning approaches utilize the complex neural network architecture to learn the representation of behaviors [75].

C. Anomaly-Based User Authentication

For the privately owned IoT devices, a large number of positive examples from the legitimate users are available while negative examples from imposters are rare. Therefore, supervised classification algorithms do not fit well to train predictive models when few negative examples are available [82]. To secure the privately owned devices, anomaly detection solutions are applied to check whether the current user is one authorized user or one imposter. Behavioral examples from imposters are referred as anomalies or outliers, and the behavioral examples from genuine users are normal. Identifying outliers or anomaly detection is referred as one-class classification. For the user authentication task, one-class

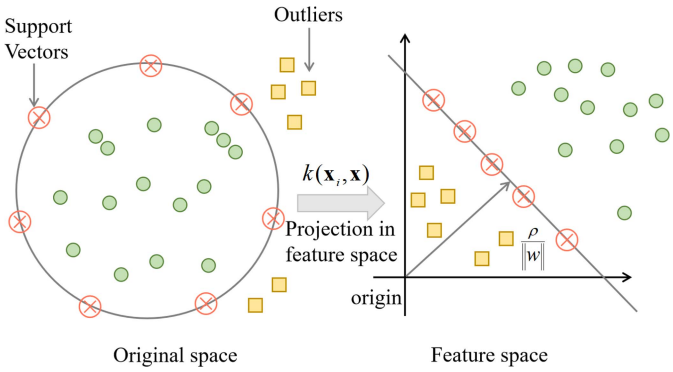

Fig. 6. General workflow of machine learning-based user identification and authentication [87].

classification algorithms, including one-class support vector machines (SVMs) [83] and isolated forest (iForest) [84] are widely adopted for imbalanced data sets with severely skewed class distributions.

One-class SVM [83] is a semisupervised classification algorithm, and aims to find a hyperplane to enclose the majority of positive examples from the origin with the maximum margin [85]. Given the training vectors $\mathbf{x}_i \in \Re^n$, the problem is formulated as follows [82]:

$$\min_{\mathbf{w},\xi,\rho} \quad \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{v^\ell}\sum_i \xi_i - \rho$$
$$\text{subject to} \quad \mathbf{w}^T \cdot \phi(\mathbf{x}_i) \geq \rho - \xi_i, \quad \xi_i \geq 0 \tag{1}$$

where $\mathbf{w}$ is the normal vector of the separating hyperplane and ξ_i are slack variables. The parameter $v \in (0, 1]$ controls the tradeoff between $\mathbf{w}$ and slack variables ξ_i. When $\mathbf{w}$ and ρ are solved by solving (1), the decision function $f(\mathbf{x}) = \text{sgn}(\sum_i \alpha_i k(\mathbf{x}_i, \mathbf{x}) - \rho)$ will be positive for majority examples, where $k(\mathbf{x}_i, \mathbf{x})$ is a kernel function [86]. In the context of behavior-based user authentication, the workflow of one-class SVM is shown in Fig. 6 to highlight how the outliers are separated from the origin. An outlier is any data instance that lies outside the support of the training data. The original high-dimensional data can be projected into one feature space via one kernel function, where the hyperplane $\mathbf{w}$ separates the training data from the origin by a maximal margin $\rho/\|\mathbf{w}\|$ (Fig. 6). Data mapped to the same side of the origin will be given a negative one-class SVM value, whereas those mapped to the side of the training data will have positive values [87].

iForest [84], [88] is one unsupervised algorithm for anomaly detection. Its main idea is based on the observation that anomalies are few in number and much different from the rest of the data [84]. Specifically, iForest constructs an ensemble of binary search trees (named iTrees) in which anomaly points are isolated closer to the root of the tree. Each node in iTree has either two children or a leaf node with no child. For one p-dimensional sample $\mathbf{x}_i \in \Re^p$ from data set $D = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_n\}$, one feature $a_i \in [1, p]$ and its split value V' are randomly selected. According to feature $V_{a_i,k}$ for each input data $\mathbf{X}_k$, $V_{a_i,k}$, which is less than V', is classified into left children set and the rest is classified into right children set. This process is repeated for the instances of left

TABLE II
SUMMARY OF WORKS IN CONTINUOUS BEHAVIORAL AUTHENTICATION AND IDENTIFICATION

Type	Work	Signal	Features	Algorithm	# of Users	Performance
Authentication	Touchstroke [29]	phone-holding behavior, keystroke dynamics	statistical features from four built-in sensors, keystroke dynamics of touch-typing timing	Bayesian Network Classifier	12	Fusion of hand movements and keystroke dynamics can improve authentication accuracy.
	HMOG [22]	accelerometer, gyroscope, and magnetometer readings when a user tapping on the screen	HMOG features (grasp resistance and grasp stability features); 11 touchscreen features; keystroke dynamics	scaled Manhattan, scaled Euclidian, One-class SVM	100	EER=7.16% for walking; EER=10.05% for sitting
	Omar et al. [31]	mouse dynamics, short-term Memory, visual scan and detection capability	statistical features of mouse dynamics	a statistical classifier based on Weighted-Sum	274	EER=2.11%
	BiLock [69]	chewing sounds	13-order Melfrequency Cepstral coefficients (MFCCs) from occlusion sounds	SVM with radial basis function (RBF) kernel	100	FAR=1.1%, FRR=5.5%
	Headbanger [66]	head movement with external rhythmic stimuli	—	dynamic time warping (DTW) distance	95	Accuracy=95.57%, FAR=4.43%
	GlassGuard [77]	touch behavior, voice features, sensor data	MFCCs, touchscreen dynamics	One-class SVM, Threshold Random Walking	32	Accuracy=93%, FAR=3%
	DeepAuth [74]	Motion Sensors	frequency domain signals	long short-term memory	47	Accuracy=96.7%
	Song et al. [78]	Keystroke dynamics	key hold time, key latency time, key duration time	isolated forest	51	Accuracy=98%
	Hong et al. [79]	Wave gesture based on accelerometer	singular value decomposition (SVD)	One-class SVM	8	Accuracy=92.83%; FPR (false positive rate)= 3.67%
Identification	George et al. [37]	Eye movement	fixation features and saccade features	Radial Basis Function Network	153	EER=2.59%
	WiFi-ID [55]	WiFi signals of walking gaits	7 time domain features and 3 frequency domain features	Sparse Approximation Classification	20	average accuracy of 93% to 77% from a group of 2 to 6 people
	WiWho [56]	WiFi signals of walking gaits	time domain and frequency domain features	decision tree	20	average accuracy of 92% to 80% from a group of 2 to 6 people
	Batchuluun et al. [75]	thermal images of walking gaits	—	convolutional neural network (CNN)	80	EER=0.77%, Accuracy=99.9%
	Mondal et al. [80]	swipe gesture	action duration, coordinate, distance, movement variability, orientation, velocity	Artificial Neural Network (ANN), Counter-Propagation ANN	41	Accuracy=95.45%
	Abo-Zahhad et al. [35]	eye blinking	amplitude, position, area, energy, slope, duration et al.	linear discriminant analysis	25	Accuracy=97.3%; EER=3.7%

and right children nodes until: 1) the incoming data set D has only one record or all data in D have the same value and 2) the tree reaches the height limit l [84], [88].

iForest is applied for user authentication due to the following reasons. First, the feature values extracted corresponding to the anomalies in original data are few and different. Second, iForest works well when handling extremely large data size and high-dimensional problems and in situations where the training set does not contain any anomalies. Finally, since iForest has linear time complexity, fast anomaly detection on resource-constrained IoT devices is crucial to report unauthorized access immediately.

D. Classification Algorithms for User Identification

User identification aims to recognize who the current user is and further determine whether the current user has the legitimate right to access the IoT devices or applications. Formally, behavior-based user identification can be formulated as follows. Given a data set $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)\}$, where $\mathbf{x}_i = [x_i^1, \ldots, x_i^m]$ is the m-dimensional feature vector of samples; $y_i \in \mathbb{C}$ is the corresponding class of one specific user; and $\mathbb{C}$ refers to the set of classes. The goal of the user identification task is to learn a mapping function that predicts the label information for one given behavior sequence with least biases. According to the adopted classification algorithms, the extant behavior-based user identification systems can be divided into two categories: 1) conventional and 2) deep learning-based solutions.

1) Conventional Classification: As shown in Table II, numerous hand-crafted features, such as statistical features of mouse dynamics, keystroke dynamics of touch-typing timing, and even frequency domain signals are proposed and a variety of supervised classification algorithms, including SVM [69], random forest [63], naïve Bayes [29], and artificial neural network [80] are employed to bridge the mapping between

feature sets and labels. However, user identification based on conventional classification algorithms relies on feature engineering, which involves computing explicit features specified by experts, resulting in algorithms designed to detect specific indictors. The hand-crafting feature is time consuming, labor intensive, and not suitable for rapidly evolving domains. Moreover, supervised learning algorithms are sensitive to the training data set. Usually, the features used in those approaches are based on samples made from the existing data set. Due to the intrauser variation of behavioral biometrics [89], these solutions are not robust when faced with the changes in user behaviors and cannot adjust the parameters accordingly.

2) Deep Learning-Based Classification: To improve the performance of identification systems, deep learning solutions are gaining popularity. Deep learning-based solutions can be divided into three categories according to the types of neural networks: 1) convolutional neural network (CNN); 2) recurrent neural network (RNN); and 3) generative adversarial network (GAN).

As shown in Fig. 7(a), a typical CNN consists of *convolutional layer*, *pooling layer*, *fully connected layer*, and *softmax layer*. In the *convolutional layer*, convolution is a linear operation that involves the multiplication of a set of weights (also referred as filter) with the input. The *convolutional layer* creates one feature map by applying the same filter on the input repetitively to summarize the presence of a specific type of features in the input. The *pooling layer* operates upon each feature map and provides an approach to downsampling feature maps by summarizing the presence of features in patches of the feature map. Two common pooling methods are average pooling and max pooling. For the *fully connected layer*, all the neurons in this layer are connected to every activation unit of the next layer operates on a flattened input where each input is connected to all the neurons. For one classification task, a *softmax layer* follows the final fully connected layer immediately to limit the output of the function into the range 0–1, which can be interpreted directly as a probability of multiple classes.

CNN is widely adopted in user authentication and identification systems to detect personal patterns from fingerprints [90] and eyes [91], [92]. CNN is also commonly applied to detect the liveness of biometrics against presentation attacks [93]–[98]. For the security of smart vehicles, Xun *et al.* designed one driver fingerprinting device for the continuous user authentication of automobiles. The driver fingerprinting device is deployed in automobiles to collect the real-time driving data from the onboard diagnostic port and uses the CNN model to extract driver behavioral characteristics from the driving data. Finally, the extracted driving features are fed to SVM for illegal driver detection [99]. To secure the smart home, Qin *et al.* [100] extracted the time and frequency features of sensors via CNN for each time slot, and fed the deep representation to RNN for user identification. Batchuluun *et al.* [75] applied CNN to identify human identity based on the walking gaits extracted from videos. Unlike gait recognition in the controlled environment, Zou *et al.* studied the user identification based on the waling gait in the wild, and proposed a hybrid model of CNN and RNN to learn robust gait feature representation, including space and time-domain features [101]

$$\begin{aligned} \mathbf{h}_t &= \tanh\left(\mathbf{W}_h \mathbf{h}_{t-1} + \mathbf{W}_y \mathbf{x}_t\right) \\ \mathbf{y}_t &= \mathbf{W}_y \mathbf{h}_t. \end{aligned} \tag{2}$$

RNN is a typical neural network to handle with sequence data, and the output of the prior state is forwarded as the input to the current state. This process can be formulated as (2), where $\mathbf{x}_t$ is the input at time t; $\mathbf{h}_t$ and $\mathbf{h}_{t-1}$ are the current state and previous state, respectively; and $\mathbf{W}_h$ and $\mathbf{W}_x$ are feedforward and recurrent weight matrices, respectively. The output $\mathbf{y}_t$ at time t is produced by combining the hidden state $\mathbf{h}_t$ with the weight matrix $\mathbf{W}_y$. However, RNNs are subject to gradient vanishing and exploding problems [102]. In addition, the training of RNN is time consuming and energy intensive [103]. To address those problems, long short-term memory (LSTM) is proposed. The architecture of one cell in LSTM is shown in Fig. 7(b). One LSTM cell consists of *forget gate*, *input gate*, and *output gate*. The *forget layer*, denoted as $\mathbf{f}_t$, controls whether the previous hide state $\mathbf{h}_{t-1}$ is forwarded to the current state $\mathbf{h}_t$, where σ is a sigmoid activation function. The *input gate* $\mathbf{i}_t$ determines to what extent new memory is added into the cells state, and an output gate $\mathbf{o}_t$ regulates how gates at the next step will be affected by the previous cell state $\mathbf{h}_{t-1}$ and current input $\mathbf{x}_t$

$$\begin{aligned} \mathbf{f}_t &= \sigma\left(\mathbf{W}_f \cdot \left[\mathbf{h}_{t-1}, \mathbf{x}_t\right] + \mathbf{b}_f\right) \\ \mathbf{i}_t &= \sigma\left(\mathbf{W}_i \cdot \left[\mathbf{h}_{t-1}, \mathbf{x}_t\right] + \mathbf{b}_i\right) \\ \widetilde{\mathbf{c}}_t &= \tanh\left(\mathbf{W}_c \cdot \left[\mathbf{h}_{t-1}, \mathbf{x}_t\right] + \mathbf{b}_c\right) \\ \mathbf{c}_t &= \mathbf{f}_t * \mathbf{c}_{t-1} + \mathbf{i}_t * \widetilde{\mathbf{c}}_t \\ \mathbf{o}_t &= \sigma\left(\mathbf{W}_o\left[\mathbf{h}_{t-1}, \mathbf{x}_t\right] + \mathbf{b}_o\right) \\ \mathbf{h}_t &= \mathbf{o}_t * \tanh(\mathbf{c}_t). \end{aligned} \tag{3}$$

Numerous user authentication and identification studies based on behavioral biometrics show that RNN and its variants are promising to process the sequential behavior data with an overwhelming performance. For example, Zhang *et al.* [104] extracted discriminate features from walking gaits monitored by smartphones and the LSTM-based model for user identification. Due to the limitation of resource-constrained devices in memory and computation, Chauhan *et al.* proposed one efficient authentication system based on breathing acoustics. They introduced model compression solutions, including weight quantization and fully connected layer factorization to reduce the complexity of LSTM, and found that compressed LSTM outperformed other baselines with smaller model size, lower inference time, and more accurate [105]. Luo *et al.* [52] used RFID for gait recognition by monitoring interruptions to RFID signals and introduced attention mechanisms in LSTM for robust user identification. Amini *et al.* [74] proposed an LSTM-based authentication framework that leveraged a user's behavior captured by motion sensors while shopping online to continuously reauthenticate the user, providing security without compromising usability. DeepAuth [106] uses unique motion patterns when users entering passwords as behavioral biometrics and learns the deep representation of motion patterns via an RNN-based model. Extensive experiments show

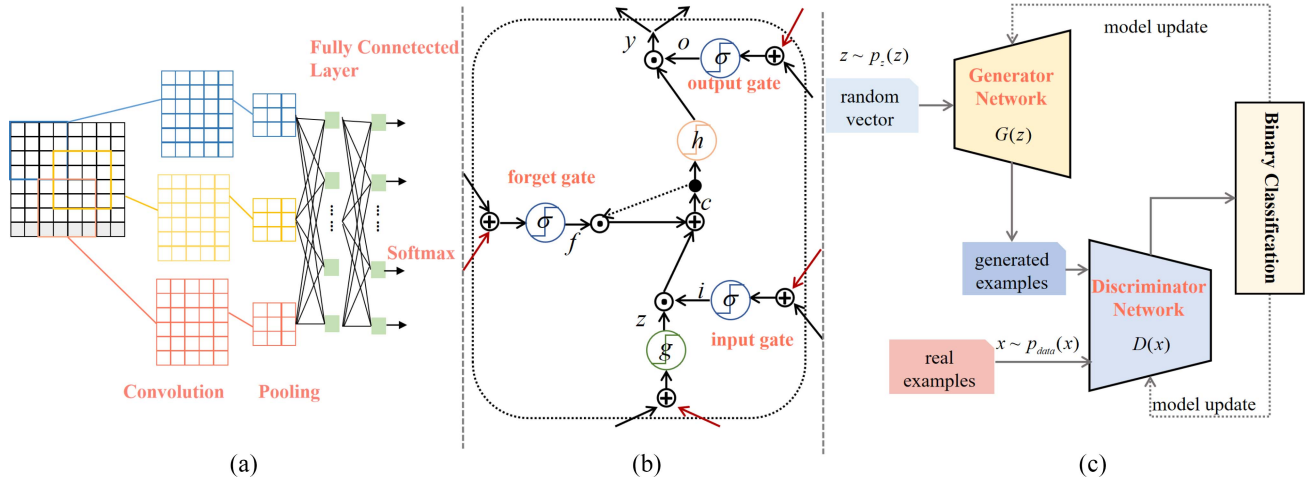

Fig. 7. Diagram of selected prevailing deep neural networks. (a) Architecture of CNN. (b) Architecture of an LSTM cell. (c) Diagram of GAN.

that DeepAuth performs well for the security of resource-constrained devices within both authentication performance and cost [106].

GAN is an unsupervised algorithm to train two competitive neural networks via a cooperative zero-sum game framework [107]. The GAN model consists of two submodels: 1) a generator model to generate new examples and 2) a discriminator model to classify whether generated examples are real data or generated examples. As shown in Fig. 7(c), the *generator network* $G(z)$ takes an random input z with probability distribution $p(z)$ and generates a sample of synthetical examples. The *discriminator network* $D(x)$ takes input either the real examples x from $p_{\text{data}}(x)$ or synthetical examples generated by $G(z)$, and attempts to predict whether the input is real or generated. The adversary learning of $G(z)$ and $D(x)$ can be represented mathematically as $\min \max E_{x\sim p_{\text{data}}(x)}[\log D(x)] + E_{z\sim p_z(z)}[\log(1 - D(G(z)))]$ [108]. GAN has been widely adopted by the adversary to generate high fidelity human biometrics, including fingerprints [109] and facial and vocal biometrics [16], [110] to bypass the authentication systems. For example, DeepMasterPrints [109] employed GAN to generate synthetic image-level fingerprints. In DeepMasterPrints, two generator networks are trained via the Wasserstein GAN algorithm based on fingerprints scanned with a capacitive sensor and a data set of inked and rolled fingerprints. The experimental results show that DeepMasterPrints is able to generate fingerprints that can easily bypass the popular commercial fingerprint matching systems.

E. Evaluation

For the evaluation of a verification system, the false acceptance rate (FAR) and the false rejection rate (FRR) are two types of errors. FAR refers the likelihood that an unauthorized user is mistakenly accepted as a legitimate user; while FRR indicates the probability that a legitimate user is incorrectly rejected as an imposter. Verification systems should avoid those two error types. To balance the two error types in a verification system, an equal error rate (EER) is introduced to predetermine the threshold value where FAR is equal to FRR. EER is a commonly accepted overall measure of system performance. The lower the EER, the higher the accuracy of the verification system.

VI. Opportunities and Challenges

CA-based behavioral traits are an emerging paradigm facing several challenges. In this section, we enumerate the opportunities and challenges from the AI-perspective.

A. Evolution of Behavioral Biometrics

The performance of CA-based behavioral biometrics could be impacted by scenarios and applications. For instance, alcohol, mood, and carrying a backpack may affect the person's gaits [111], [112]. Pupil size and eye movement dynamics vary with individuals' physical status, such as stress or fatigue [113]. In addition, touch patterns are sensitive to screen size and target applications. Obviously, user behaviors can be impaired by many facets, including mood, health, and alcohol. How to rule out the impacts of situations on user behaviors in the CA systems is crucial for the robustness of behavior-based authentication systems.

User behaviors also change over time. For example, Galbally *et al.* analyzed the effects of age and aging on fingerprints, and found that fingerprint quality decreased linearly with age for elders [114]. The touchscreen typing patterns are dynamic and impacted by health status [115]. However, what kinds of behavioral patterns can be used as behavioral biometrics for CA have not been studied yet. In addition, the extant behavior-based systems are static and cannot adjust themselves with the evolution of user behaviors. Therefore, fundamental studies should be conducted to evaluate the impacts of dynamics of behavioral traits on CA and propose dynamic solutions that fit well with the changes in behavior evolution.

B. Sparsity of Behavioral Biometrics

The success of machine learning roots in tapping into the large-scale training data set to learn the comprehensive representation of the rapid advances of AI-based methods. A large number of users are becoming multidevice users by interacting with more than one IoT devices [116]. The one-to-many mapping relationship between users and IoT devices poses great

challenges for CA-based on behavioral biometrics. In addition, the sufficient training data set is not readily available, and the trained models based on the skewed distribution data sets are not robust [117].

To address the sparsity of behavioral biometrics, one/zero-shot learning aims to build one classifier from one or only a few training samples [118]. The nature of one/zero-shot learning is suitable for handling the sparsity problem of behavioral biometrics. Meanwhile, as user preferences of multiple IoT devices are significantly different, the interaction records collected from different IoT devices are different. For example, the keystroke rhythm can be captured when interacting with PC; and the touchscreen dynamics are collected via the mobile phone or tablets. To handle the imbalanced classification problem and smoothly conduct the user authentication on heterogeneous IoT devices, transfer learning [119], [120] is promising to transfer the trained model based on behavioral records collected on one source device to the target device. In addition, to address the sparsity of labeled data, AutoTune uses the wireless identifier as a supervisory label and learns the association between facial images and wireless identifier [121].

C. Deep Learning on Resource-Constrained IoT Devices

In general, IoT devices, including smartphones and wearables are resource-constrained with limited memory, power supply, and computing capability [122]. Although deep learning algorithms achieved state-of-the-art performance, deep learning models are becoming extremely complex with millions of hyperparameters [123], time-consuming training, significant energy strains [103], [124], and do not work well on resource-constrained IoT devices by *debilitating levels of system overhead* [122]. To execute deep learning-based models on resource-constrained devices, the following two types of solutions are proposed.

Model compression of deep learning has become a significant problem. Methods to reduce the complexity of deep neural networks include tensor decomposition [125], [126], pruning [124], [127], and parameter sharing [128]–[130]. Tensor decomposition reduces model complexity by expressing a higher order tensor with a sequence of linear operations on the matrix singular value decomposition [126]. Neural pruning eliminates the less important connections in one pretrained model to reduce the computational cost by compressing hyperparameters [131] or multiobjective optimization based on accuracy, latency, and energy [132]. Parameter sharing [133] is mainly applied in convolutional layers to reduce the size of parameters by the assumption that the input going to be processed by the network is decomposable into a set of local regions with the same nature and thus each of them can be processed with the same set of transformations [130], [133].

Edge computing is shedding lights on mobile devices-oriented deep learning [134], [135]. Via edge intelligence, the complex and energy-intensive deep neural networks can be partitioned into tiny subtasks, and be distributively executed on neighboring devices or edges [136]–[138]. For example, Kang *et al.* [136] designed a lightweight scheduler to automatically balance the computational offload between mobile devices and servers by partitioning the neural network layers. Xu *et al.* [138] proposed DeepWear to optimize the energy consumption of wearables by offloading the deep learning tasks among mobile devices.

Although there are many pilot studies about deep learning on resource-constrained IoT devices, they are evaluated in the laboratory environments. As a result, the performance is not generalizable to other contexts (e.g., smart home) due to the extreme heterogeneity of IoT environments [139]. In addition, CA on resource-constrained IoT devices is very sensitive to latency. A novel framework is essential to understand the tradeoff among accuracy, critical latency, and efficiency [140].

D. Emerging Malicious Attacks

CA based on behavioral biometrics cannot be easily attacked by a random attacker. However, the nature of CA based on behavioral biometrics does not imply that the CA systems are secure.

First, most authentication systems-based behavioral biometrics are prototypes evaluated in a constrained laboratory environment with limited participants. Comprehensive evaluations with a large number of participants are needed to investigate the performance of existing behavior-based authentication systems when faced with potential attacks.

Second, IoT systems are faced with numerous security threats on physical, protocol, communication, and application layers [27]. For example, in the low-power wireless network, the energy depletion attack can drain the batteries of devices rapidly by forcing sensors or actuators to execute energy-intensive tasks. Consequently, the entire network could fail due to battery exhaustion [141]. For the electrical vehicles, batteries could be attacked (e.g., draining energy) to reduce driving range and increase driving range anxiety. Kang and Shen [142] proposed a battery authentication method based on user habits to identify users that share a vehicle. In addition, the thermal attacks rely on the heat transferred from users to interactive devices, and exploit heat traces in the wake of user interaction with devices to uncover the entered credentials [11]. On the communication layer, the heterogeneous communication protocols are subject to attacks, such as eavesdropping, sinkhole, hello flood, and collision [143], [144]. Voice assistants, such as Google Assistant, Amazon Alexa, Facebook Portal, and Apple Siri are vulnerable to signal injection attacks on microphones based on the photoacoustic effect across large distances and through glass windows.[1]

Third, AI is widely applied by imposters to hack the authentication and identification systems. For example, AI has been maliciously used by the imposters to infer the password. Snoopy demonstrates that it is possible to infer password entered on mobile devices by monitoring both accelerometers and gyroscopes [145]. Snoopy may fool the potential users as a harmless app to continuously monitor the motion sensors when users are taping the passwords, and uses bidirectional RNN for most commonly used password inference and encoder–decoder architecture with RNN models for universal password interference [145]. Another prevailing example is

[1]https://lightcommands.com/

the usage of AI by the adversary to reconstruct the biometrics (i.e., replay attack). Deep generative models (DGMs), such as GANs and variational autoencoder (VAE) have been widely adopted to generate high fidelity human biometrics, including fingerprints [109] and facial and vocal biometrics [16], [110] to bypass the authentication systems.

E. Fusion of Behavioral Biometrics

Since prior solutions are insufficient to effectively provide secure protection in a broad range of IoT scenarios, multifactor authentication (MFA) provides multiple layers of security to protect IoT devices against potential attacks through the fusion of behavioral biometrics [2].

The early form of MFA integrated multiple authentication schemes sequentially. For example, Hu *et al.* [146] proposed a secure data backup scheme by integrating password and biometrics to overcome the potential attacks. Multiview representation learning for user authentication has emerged as a viable approach to process such data. This paradigm of machine learning aims to fuse multiple views (i.e., feature sets) to improve the performance [147], [148]. Multiview representation learning can be categorized into two groups: 1) multimodal methods and 2) multiview methods.

Multimodal solutions extract features from heterogeneous biometrics to build one classifier based on the ensembled features. Kim *et al.* [149] designed one multimodal authentication system by fusing features obtained from face, teeth, and voice modalities to secure mobile devices. Crawford *et al.* utilized the keystroke dynamics and speaker verification to enhance the authentication performance on mobile devices with a 67% reduction of explicit authentication [150]. EchoPrint emits nearly inaudible sound signals from the earpiece speaker to illuminate the user's face. The extracted acoustic features from the echoes are combined with visual facial landmarks from the frontal camera to authenticate the user [151]. Gomi *et al.* [152] integrated physiological biometrics, behavioral traits, and online activities, including search, shopping, and Web browsing for user authentication. Kumar *et al.* utilized LSTM to model the motion sensors and adopted CNN to extract gait patterns from video, respectively, and a Gray wolf optimizer has been used to optimize the parameters during fusion [153]. VAuth collects the body vibrations of the user and matches it with the speech signal received by the voice assistant's microphone. By fusing multimodal data, VAuth shows robust performance against potential attacks, including replay attacks, mangled voice attacks, or impersonation attacks [154]. To against the replay attacks, REVOLT exploits the spectral differences between original and replayed voice signals, and combines the breathing rate extracted from the WiFi signal while speaking to detect the liveness [155].

Multiview-based solutions for user authentication mainly extracted fine-grained feature maps from multiview images. For example, Li *et al.* [156] employed the multiview deep representation learning to recognize one million celebrities from their face images captured in the real world. On the other hand, multiview learning can be applied for user authentication to address the incompletion of obtained biometrics by fusing multiple views [157].

In addition, contextual information can enhance the performance of user authentication. Hintze *et al.* [158] introduced dynamic factors, such as day and time, and location together with multimodal biometrics to adjust the authentication scheme accordingly. Wójtowicz and Joachimiak [159] presented one context-based biometric authentication model, which chooses the appropriate authentication method dynamically according to the interaction form.

F. Cross-Device Continuous Authentication

Increasingly, a large number of users are becoming multidevice users by interacting with multiple smart devices [116], [160]. For instance, more than 70% of online users access the Internet across multiple devices. 90% use multiple screens sequentially to accomplish a task over time.[2] The complexity of multidevice–multiuser interaction presents significant challenges for cross-device CA. Prior studies assume the one-to-one mapping between user and device and mainly focus on the user authentication in the single-device scenario. However, the relationships between users and devices in multiuser–multidevice scenarios are many to many. Due to the heterogeneity of devices, transferring one pretrained user identification model from the source domain to the target domain is rarely studied. There are significant differences in interaction modality among heterogeneous devices. For devices with similar interaction modalities, transfer learning is one promising solution [116]. In addition, the co-location information of devices is useful for cross-device authentication. Hintze *et al.* [160] proposed one multimodal and cross-device authentication system based on behavioral and physiological biometrics (e.g., gait, voice, face, and keystroke dynamics) to reduce the manual burden of user verification according to the context, such as location, time of day, and nearby devices.

G. Privacy Concerns

Behavior-based authentication systems are subject to privacy concerns, especially when they are adopted for personalized services. To address privacy concerns, many privacy protection solutions are proposed to secure crucial information in different authentication systems [161].

For the standalone authentication systems, although users are able to control the standalone client, they are likely to subject to exposing data to unauthorized third parties [161]. To secure the sensitive data in standalone systems, privacy impact assessment and surveillance impact assessment should be enforced to ensure conformance with legal and regulatory requirements [161]. Moreover, anonymity and encryption are promising to protect against data exposure in privately owned devices. For the centralized authentication systems, they may suffer from data exposure in transit and adversary attacks in computation models by carefully crafted adversarial samples [162]. Blockchain keeps the sensitive data private such that others cannot trace and infer sensitive data stored in the block [163]. However, the centralized authentication systems

[2]https://www.readinbrief.com/multi-device-content-consumption-statistics-trends/

may suffer from the bottle problem due to the limitation of a single centralized server [7].

For the distributed authentication systems, blockchain and federated learning [164], [165] show a great potential to provide privacy-preserving authentication in collaborative applications. Federated learning is a machine learning technique that trains an algorithm across multiple decentralized edge devices or servers holding local data samples, without exchanging their data samples. This approach contrasts with traditional centralized machine learning techniques where all data samples are uploaded to one server, as well as to more classical decentralized approaches which assume that local data samples are identically distributed [164]–[166]. The nature of federated learning not only can prevent data sharing among devices but also avoid the enormous communication costs. In addition, other privacy-preserving machine learning approaches, such as multiparty computation (MPC) and homomorphic encryption are getting more attraction recently. For intrusted participants, the MPC is able to calculate a joint function in a decentralized network on the premise of ensuring privacy and independence of input [167].

VII. Conclusion

In the IoT era, user authentication and identification are critical to ensure the security of connected things and the customization of passive services. However, conventional identification methods suffer from several key drawbacks, including discreteness, obtrusiveness, and vulnerability. In this article, we propose the CA based on behavioral biometrics, characterize the key features of CA based on user behaviors (e.g., invulnerability, continuity, unobtrusiveness, and convenience), and summarize the existing CA solutions from sensing and computing. Based on this taxonomy, we discuss the challenges and open issues from the perspective of AI.

References

[1] D. J. Cook, J. C. Augusto, and V. R. Jakkula, "Ambient intelligence: Technologies, applications, and opportunities," *Pervasive Mobile Comput.*, vol. 5, no. 4, pp. 277–298, 2009.

[2] A. Ometov, V. Petrov, S. Bezzateev, S. Andreev, Y. Koucheryavy, and M. Gerla, "Challenges of multi-factor authentication for securing advanced IoT applications," *IEEE Netw.*, vol. 33, no. 2, pp. 82–88, Mar./Apr. 2019.

[3] S. Gupta, A. Buriro, and B. Crispo, "DriverAuth: A risk-based multimodal biometric-based driver authentication scheme for ride-sharing platforms," *Comput. Security*, vol. 83, pp. 122–139, Jun. 2019.

[4] Y. Sun, B. Wang, S. Li, Z. Sun, H. M. Nguyen, and T. Q. Duong, "Manipulation with domino effect for cache- and buffer-enabled social IIoT: Preserving stability in tripartite graphs," *IEEE Trans. Ind. Informat.*, vol. 16, no. 8, pp. 5389–5400, Aug. 2020.

[5] A. K. Jain, A. Ross, and S. Pankanti, "Biometrics: A tool for information security," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 2, pp. 125–143, Jun. 2006.

[6] L. Tang, Z. Duan, Y. Zhu, J. Ma, and Z. Liu, "Recommendation for ridesharing groups through destination prediction on trajectory data," *IEEE Trans. Intell. Transp. Syst.*, early access, Dec. 27, 2019, doi: 10.1109/TITS.2019.2961170.

[7] H. Liu, P. Zhang, G. Pu, T. Yang, S. Maharjan, and Y. Zhang, "Blockchain empowered cooperative authentication with data traceability in vehicular edge computing," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4221–4232, Apr. 2020.

[8] T. Kwon and J. Hong, "Analysis and improvement of a pin-entry method resilient to shoulder-surfing and recording attacks," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 2, pp. 278–292, Feb. 2015.

[9] M. Čagalj, T. Perković, and M. Bugarić, "Timing attacks on cognitive authentication schemes," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 3, pp. 584–596, Mar. 2015.

[10] T. Chen, M. Farcasin, and E. Chan-Tin, "Smartphone passcode prediction," *IET Inf. Security*, vol. 12, no. 5, pp. 431–437, 2018.

[11] Y. Abdelrahman, M. Khamis, S. Schneegass, and F. Alt, "Stay cool! Understanding thermal attacks on mobile-based user authentication," in *Proc. ACM Conf. Human Factors Comput. Syst. (CHI)*, 2017, pp. 3751–3763.

[12] D. Li, X.-P. Zhang, M. Hu, G. Zhai, and X. Yang, "Physical password breaking via thermal sequence analysis," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 5, pp. 1142–1154, May 2019.

[13] A. J. Aviv, K. Gibson, E. Mossop, M. Blaze, and J. M. Smith, "Smudge attacks on smartphone touch screens," in *Proc. 4th USENIX Conf. Offensive Technol. (WOOT)*, Berkeley, CA, USA, 2010, pp. 1–7.

[14] S. Schneegass, F. Steimle, A. Bulling, F. Alt, and A. Schmidt, "SmudgeSafe: Geometric image transformations for smudge-resistant user authentication," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. (UbiComp)*, 2014, pp. 775–786.

[15] S. Kim, H. Lee, and T. Kwon, "Poster: Rethinking fingerprint identification on smartphones," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security (CCS)*, 2017, pp. 2515–2517.

[16] A. Roy, N. Memon, and A. Ross, "MasterPrint: Exploring the vulnerability of partial fingerprint-based authentication systems," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 9, pp. 2013–2025, Sep. 2017.

[17] Y. Ren, Z. Fang, D. Liu, and C. Chen, "Replay attack detection based on distortion by loudspeaker for voice authentication," *Multimedia Tools Appl.*, vol. 78, pp. 8383–8396, Nov. 2019.

[18] Y. Wang, W. Cai, T. Gu, W. Shao, Y. Li, and Y. Yu, "Secure your voice: An oral airflow-based continuous liveness detection for voice assistants," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 3, no. 4, p. 157, Dec. 2019. [Online]. Available: https://doi.org/10.1145/3369811

[19] M. Gil, P. Giner, and V. Pelechano, "Personalization for unobtrusive service interaction," *Pers. Ubiquitous Comput.*, vol. 16, no. 5, pp. 543–561, Jun. 2012.

[20] Y. Sui, X. Zou, E. Y. Du, and F. Li, "Design and analysis of a highly user-friendly, secure, privacy-preserving, and revocable authentication method," *IEEE Trans. Comput.*, vol. 63, no. 4, pp. 902–916, Apr. 2014.

[21] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song, "Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 1, pp. 136–148, Jan. 2013.

[22] Z. Sitová *et al.*, "HMOG: New behavioral biometric features for continuous authentication of smartphone users," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 5, pp. 877–892, May 2016.

[23] Y.-L. Zheng *et al.*, "Unobtrusive sensing and wearable devices for health informatics," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1538–1554, May 2014.

[24] M. Weiser, "The computer for the 21st century," *IEEE Pervasive Comput.*, vol. 1, no. 1, pp. 19–25, Jul. 2002.

[25] W. Meng, D. S. Wong, S. Furnell, and J. Zhou, "Surveying the development of biometric user authentication on mobile phones," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1268–1293, 3rd Quart., 2015.

[26] N. Sae-Bae, J. Wu, N. Memon, J. Konrad, and P. Ishwar, "Emerging NUI-based methods for user authentication: A new taxonomy and survey," *IEEE Trans. Biometr. Behav. Identity Sci.*, vol. 1, no. 1, pp. 5–31, Jan. 2019.

[27] A. Mosenia and N. K. Jha, "A comprehensive study of security of Internet-of-Things," *IEEE Trans. Emerg. Topics Comput.*, vol. 5, no. 4, pp. 586–602, Oct. 2017.

[28] C. X. Lu, H. Wen, S. Wang, A. Markham, and N. Trigoni, "SCAN: Learning speaker identity from noisy sensor data," in *Proc. 16th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, 2017, pp. 67–78.

[29] A. Buriro, B. Crispo, F. Del Frari, and K. Wrona, "Touchstroke: Smartphone user authentication based on touch-typing biometrics," in *New Trends in Image Analysis and Processing*, V. Murino, E. Puppo, D. Sona, M. Cristani, and C. Sansone, Eds. Cham, Switzerland: Springer, 2015, pp. 27–34.

[30] A. J. Aviv, B. Sapp, M. Blaze, and J. M. Smith, "Practicality of accelerometer side channels on smartphones," in *Proc. 28th Annu. Comput. Security Appl. Conf. (ACSAC)*, 2012, pp. 41–50.

[31] O. Hamdy and I. Traoré, "Homogeneous physio-behavioral visual and mouse-based biometric," *ACM Trans. Comput. Human Interact.*, vol. 18, no. 3, p. 12, Aug. 2011.

[32] N. Zheng, A. Paloski, and H. Wang, "An efficient user verification system using angle-based mouse movement biometrics," *ACM Trans. Inf. Syst. Security*, vol. 18, no. 3, pp. 1–27, Apr. 2016.

[33] J. Roh, S. Lee, and S. Kim, "Keystroke dynamics for authentication in smartphone," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2016, pp. 1155–1159.

[34] L. Lu and Y. Liu, "Safeguard: User reauthentication on smartphones via behavioral biometrics," *IEEE Trans. Comput. Soc. Syst.*, vol. 2, no. 3, pp. 53–64, Sep. 2015.

[35] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "A novel biometric approach for human identification and verification using eye blinking signal," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 876–880, Jul. 2015.

[36] Y. Zhang, W. Hu, W. Xu, C. T. Chou, and J. Hu, "Continuous authentication using eye movement response of implicit visual stimuli," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 1, no. 4, pp. 1–22, Jan. 2018.

[37] A. George and A. Routray, "A score level fusion method for eye movement biometrics," *Pattern Recognit. Lett.*, vol. 82, pp. 207–215, Oct. 2016.

[38] S. Eberz, K. B. Rasmussen, V. Lenders, and I. Martinovic, "Looks like eve: Exposing insider threats using eye movement biometrics," *ACM Trans. Privacy Security*, vol. 19, no. 1, pp. 1–31, Jun. 2016.

[39] Z. Wu, Y. Huang, L. Wang, X. Wang, and T. Tan, "A comprehensive study on cross-view gait based human identification with deep CNNs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 2, pp. 209–226, Feb. 2017.

[40] J. P. Singh, S. Jain, S. Arora, and U. P. Singh, "Vision-based gait recognition: A survey," *IEEE Access*, vol. 6, pp. 70497–70527, 2018.

[41] Y. Liang, X. Zhou, Z. Yu, and B. Guo, "Energy-efficient motion related activity recognition on mobile devices for pervasive healthcare," *Mobile Netw. Appl.*, vol. 19, no. 3, pp. 303–317, Jun. 2014.

[42] M. U. B. Altaf, T. Butko, and B. Juang, "Acoustic gaits: Gait analysis with footstep sounds," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 8, pp. 2001–2011, Aug. 2015.

[43] H. Ma and W. Liao, "Human gait modeling and analysis using a semi-Markov process with ground reaction forces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 6, pp. 597–607, Jun. 2017.

[44] P. Casale, O. Pujol, and P. Radeva, "Personalization and user verification in wearable systems using biometric walking patterns," *Pers. Ubiquitous Comput.*, vol. 16, no. 5, pp. 563–580, Jun. 2012.

[45] D. Ma, G. Lan, W. Xu, M. Hassan, and W. Hu, "Unobtrusive user verification using piezoelectric energy harvesting," in *Proc. 14th EAI Int. Conf. Mobile Ubiquitous Syst. Comput. Netw. Services (MobiQuitous)*, 2017, pp. 541–542.

[46] D. Baek, P. Musale, and J. Ryoo, "Walk to show your identity: Gait-based seamless user authentication framework using deep neural network," in *Proc. 5th ACM Workshop Wearable Syst. Appl. (WearSys)*, 2019, pp. 53–58.

[47] G. Cola, M. Avvenuti, F. Musso, and A. Vecchio, "Gait-based authentication using a wrist-worn device," in *Proc. 13th Int. Conf. Mobile Ubiquitous Syst. Comput. Netw. Services (MOBIQUITOUS)*, 2016, pp. 208–217.

[48] M. D. Marsico and A. Mecca, "A survey on gait recognition via wearable sensors," *ACM Comput. Surveys*, vol. 52, no. 4, pp. 1–39, Aug. 2019.

[49] Y. Ren, Y. Chen, M. C. Chuah, and J. Yang, "User verification leveraging gait recognition for smartphone enabled mobile healthcare systems," *IEEE Trans. Mobile Comput.*, vol. 14, no. 9, pp. 1961–1974, Sep. 2015.

[50] F. Lin, A. Wang, Y. Zhuang, M. R. Tomita, and W. Xu, "Smart insole: A wearable sensor device for unobtrusive gait monitoring in daily life," *IEEE Trans. Ind. Informat.*, vol. 12, no. 6, pp. 2281–2291, Dec. 2016.

[51] M. Muaaz and R. Mayrhofer, "Smartphone-based gait recognition: From authentication to imitation," *IEEE Trans. Mobile Comput.*, vol. 16, no. 11, pp. 3209–3221, Nov. 2017.

[52] C. Luo *et al.*, "Gait recognition as a service for unobtrusive user identification in smart spaces," *ACM Trans. Internet Things*, vol. 1, no. 1, pp. 1–21, Mar. 2020.

[53] F. Xiao, Z. Guo, Y. Ni, X. Xie, S. Maharjan, and Y. Zhang, "Artificial intelligence empowered mobile sensing for human flow detection," *IEEE Netw.*, vol. 33, no. 1, pp. 78–83, Jan./Feb. 2019.

[54] Z. Wang, B. Guo, Z. Yu, and X. Zhou, "Wi-Fi CSI-based behavior recognition: From signals and actions to activities," *IEEE Commun. Mag.*, vol. 56, no. 5, pp. 109–115, May 2018.

[55] J. Zhang, B. Wei, W. Hu, S. S. Kanhere, and A. Tan, "Human identification using WiFi signal," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops)*, Mar. 2016, pp. 1–2.

[56] Y. Zeng, P. H. Pathak, and P. Mohapatra, "WiWho: WiFi-based person identification in smart spaces," in *Proc. 15th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2016, pp. 1–12.

[57] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using WiFi signals," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. (UbiComp)*, 2016, pp. 363–373.

[58] T. Xin *et al.*, "FreeSense: A robust approach for indoor human detection using Wi-Fi signals," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, pp. 1–23, Sep. 2018.

[59] D. Gafurov, E. Snekkenes, and P. Bours, "Gait authentication and identification using wearable accelerometer sensor," in *Proc. IEEE Workshop Autom. Identification Adv. Technol.*, Jun. 2007, pp. 220–225.

[60] R. Cham and M. S. Redfern, "Changes in gait when anticipating slippery floors," *Gait Posture*, vol. 15, no. 2, pp. 159–171, 2002.

[61] K. Singhal and J. B. Casebolt, "Chapter 8—Aging and gait," in *Nutrition and Functional Foods for Healthy Aging*, R. R. Watson, Ed. London, U.K.: Academic, 2017, pp. 65–74.

[62] J. Liang, Y. Cao, C. Zhang, S. Chang, K. Bai, and Z. Xu, "Additive adversarial learning for unbiased authentication," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2019, pp. 11420–11429.

[63] K. Diederichs, A. Qiu, and G. Shaker, "Wireless biometric individual identification utilizing millimeter waves," *IEEE Sens. Lett.*, vol. 1, no. 1, pp. 1–4, Feb. 2017.

[64] N. Zhao *et al.*, "Authentication in millimeter-wave body-centric networks through wireless channel characterization," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6616–6623, Dec. 2017.

[65] K. Matsuo, F. Okumura, M. Hashimoto, S. Sakazawa, and Y. Hatori, "Arm swing identification method with template update for long term stability," in *Advances in Biometrics*, S.-W. Lee and S. Z. Li, Eds. Berlin, Germany: Springer, 2007, pp. 211–221.

[66] S. Li, A. Ashok, Y. Zhang, C. Xu, J. Lindqvist, and M. Gruteser, "Whose move is it anyway? Authenticating smart wearable devices using unique head movement patterns," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Mar. 2016, pp. 1–9.

[67] L. Lu *et al.*, "Lip reading-based user authentication through acoustic sensing on smartphones," *IEEE/ACM Trans. Netw.*, vol. 27, no. 1, pp. 447–460, Feb. 2019.

[68] D. Lu, D. Huang, Y. Deng, and A. Alshamrani, "Multifactor user authentication with in-air-handwriting and hand geometry," in *Proc. Int. Conf. Biometr. (ICB)*, Feb. 2018, pp. 255–262.

[69] Y. Zou, M. Zhao, Z. Zhou, J. Lin, M. Li, and K. Wu, "BiLock: User authentication via dental occlusion biometrics," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 3, pp. 1–20, Sep. 2018.

[70] T. Rahman *et al.*, "BodyBeat: A mobile system for sensing non-speech body sounds," in *Proc. 12th Annu. Int. Conf. Mobile Syst. Appl. Services (MobiSys)*, 2014, pp. 2–13.

[71] J. Chauhan, S. Seneviratne, Y. Hu, A. Misra, A. Seneviratne, and Y. Lee, "Breathing-based authentication on resource-constrained IoT devices using recurrent neural networks," *Computer*, vol. 51, no. 5, pp. 60–67, May 2018.

[72] J. Liu, Y. Dong, Y. Chen, Y. Wang, and T. Zhao, "Leveraging breathing for continuous user authentication," in *Proc. ACM 24th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2018, pp. 786–788. [Online]. Available: http://doi.acm.org/10.1145/3241539.3267743

[73] A. Alzubaidi and J. Kalita, "Authentication of smartphone users using behavioral biometrics," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1998–2026, 3rd Quart., 2016.

[74] S. Amini, V. Noroozi, A. Pande, S. Gupte, P. S. Yu, and C. Kanich, "DeepAuth: A framework for continuous user re-authentication in mobile apps," in *Proc. 27th ACM Int. Conf. Inf. Knowl. Manag. (CIKM)*, 2018, pp. 2027–2035.

[75] G. Batchuluun, R. A. Naqvi, W. Kim, and K. R. Park, "Body-movement-based human identification using convolutional neural network," *Expert Syst. Appl.*, vol. 101, pp. 56–77, Jul. 2018.

[76] Y. Yang, B. Guo, Z. Wang, M. Li, Z. Yu, and X. Zhou, "BehaveSense: Continuous authentication for security-sensitive mobile apps using behavioral biometrics," *Ad Hoc Netw.*, vol. 84, pp. 9–18, Mar. 2019.

[77] G. Peng, D. T. Nguyen, G. Zhou, and S. Wang, "Poster: A continuous and noninvasive user authentication system for Google glass," in *Proc. 13th Annu. Int. Conf. Mobile Syst. Appl. Services (MobiSys)*, 2015, pp. 487–487.

[78] K. Song, Y. Zhou, H. Liu, and N. Zhu, "Isolated forest in keystroke dynamics-based authentication: Only normal instances available for training," in *Proc. 2nd IEEE Int. Conf. Comput. Intell. Appl. (ICCIA)*, 2017, pp. 63–67.

[79] F. Hong, M. Wei, S. You, Y. Feng, and Z. Guo, "Waving authentication: Your smartphone authenticate you on motion gesture," in *Proc. 33rd Annu. ACM Conf. Extended Abstracts Human Factors Comput. Syst. (CHI EA)*, 2015, pp. 263–266.

[80] S. Mondal and P. Bours, "Continuous authentication and identification for mobile devices: Combining security and forensics," in *Proc. IEEE Int. Workshop Inf. Forensics Security (WIFS)*, Nov. 2015, pp. 1–6.

[81] Z. Cai, C. Shen, and X. Guan, "Mitigating behavioral variability for mouse dynamics: A dimensionality-reduction-based approach," *IEEE Trans. Human–Mach. Syst.*, vol. 44, no. 2, pp. 244–255, Apr. 2014.

[82] L. M. Manevitz and M. Yousef, "One-class SVMs for document classification," *J. Mach. Learn. Res.*, vol. 2, pp. 139–154, Mar. 2002.

[83] B. Schölkopf, R. Williamson, A. Smola, J. Shawe-Taylor, and J. Platt, "Support vector method for novelty detection," in *Proc. 12th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, Cambridge, MA, USA, 1999, pp. 582–588.

[84] F. T. Liu, K. M. Ting, and Z. Zhou, "Isolation forest," in *Proc. 8th IEEE Int. Conf. Data Min.*, 2008, pp. 413–422.

[85] D. Shin and S. Kim, "Nearest mean classification via one-class SVM," in *Proc. Int. Joint Conf. Comput. Sci. Optim.*, vol. 1, 2009, pp. 593–596.

[86] H. Lukashevich, S. Nowak, and P. Dunker, "Using one-class SVM outliers detection for verification of collaboratively tagged image training sets," in *Proc. IEEE Int. Conf. Multimedia Expo*, 2009, pp. 682–685.

[87] Y. Guerbai, Y. Chibani, and N. Abbas, "One-class versus bi-class SVM classifier for off-line signature verification," in *Proc. Int. Conf. Multimedia Comput. Syst.*, 2012, pp. 206–210.

[88] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation-based anomaly detection," *ACM Trans. Knowl. Discovery Data*, vol. 6, no. 1, pp. 1–39, Mar. 2012.

[89] L. M. Mayron, "Biometric authentication on mobile devices," *IEEE Security Privacy*, vol. 13, no. 3, pp. 70–73, May/Jun. 2015.

[90] R. Das, E. Piciucco, E. Maiorana, and P. Campisi, "Convolutional neural network for finger-vein-based biometric identification," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 2, pp. 360–373, Feb. 2019.

[91] Z. Zhao and A. Kumar, "Improving periocular recognition by explicit attention to critical regions in deep neural network," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 12, pp. 2937–2952, Dec. 2018.

[92] Q. Zhang, H. Li, Z. Sun, and T. Tan, "Deep feature fusion for iris and periocular biometrics on mobile devices," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 11, pp. 2897–2912, Nov. 2018.

[93] E. Park, X. Cui, T. H. B. Nguyen, and H. Kim, "Presentation attack detection using a tiny fully convolutional network," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 11, pp. 3016–3025, Nov. 2019.

[94] R. F. Nogueira, R. de Alencar Lotufo, and R. C. Machado, "Fingerprint liveness detection using convolutional neural networks," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 6, pp. 1206–1213, Jun. 2016.

[95] W. Kang, H. Liu, W. Luo, and F. Deng, "Study of a full-view 3D finger vein verification technique," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 1175–1189, 2020.

[96] A. Czajka, "Pupil dynamics for iris liveness detection," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 4, pp. 726–735, Apr. 2015.

[97] O. V. Komogortsev, A. Karpov, and C. D. Holland, "Attack of mechanical replicas: Liveness detection with eye movements," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 4, pp. 716–725, Apr. 2015.

[98] C. Yuan *et al.*, "Fingerprint liveness detection using an improved CNN with image scale equalization," *IEEE Access*, vol. 7, pp. 26953–26966, 2019.

[99] Y. Xun, J. Liu, N. Kato, Y. Fang, and Y. Zhang, "Automobile driver fingerprinting: A new machine learning based authentication scheme," *IEEE Trans. Ind. Informat.*, vol. 16, no. 2, pp. 1417–1426, Feb. 2020.

[100] Z. Qin *et al.*, "Learning-aided user identification using smartphone sensors for smart homes," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7760–7772, Oct. 2019.

[101] Q. Zou, Y. Wang, Q. Wang, Y. Zhao, and Q. Li, "Deep learning-based gait recognition using smartphones in the wild," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3197–3212, Apr. 2020.

[102] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Proc. 30th Int. Conf. Mach. Learn. (ICML)*, vol. 28, 2013, pp. 1310–1318.

[103] E. Strubell, A. Ganesh, and A. McCallum, "Energy and policy considerations for deep learning in NLP," in *Proc. 57th Annu. Meeting Assoc. Comput. Linguist.*, Jul. 2019, pp. 3645–3650.

[104] H. Zhang, J. Liu, K. Li, H. Tan, and G. Wang, "Gait learning based authentication for intelligent things," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4450–4459, Apr. 2020.

[105] J. Chauhan, J. Rajasegaran, S. Seneviratne, A. Misra, A. Seneviratne, and Y. Lee, "Performance characterization of deep learning models for breathing-based authentication on resource-constrained devices," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 4, p. 158, Dec. 2018.

[106] C. X. Lu *et al.*, "DeepAuth: *In-situ* authentication for smartwatches via deeply learned behavioural biometrics," in *Proc. ACM Int. Symp. Wearable Comput. (ISWC)*, 2018, pp. 204–207.

[107] I. J. Goodfellow *et al.*, "Generative adversarial nets," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, vol. 2, 2014, pp. 2672–2680.

[108] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 53–65, Jan. 2018.

[109] P. Bontrager, A. Roy, J. Togelius, N. Memon, and A. Ross, "DeepMasterPrints: Generating MasterPrints for dictionary attacks via latent variable evolution," in *Proc. IEEE 9th Int. Conf. Biometr. Theory Appl. Syst. (BTAS)*, 2018, pp. 1–9.

[110] A. van den Oord *et al.*, "WaveNet: A generative model for raw audio," Sep. 2016. [Online]. Available: arXiv:1609.03499.

[111] H. N. Ahmad and T. M. Barbosa, "The effects of backpack carriage on gait kinematics and kinetics of schoolchildren," *Sci. Rep.*, vol. 9, no. 3364, pp. 86–95, 2019.

[112] G. F. D. Greenstein, "Gait and balance deficits in chronic alcoholics: No improvement from 10 weeks through 1 year abstinence," *Alcoholism Clin. Exp. Res.*, vol. 37, no. 1, pp. 89–95, 2012.

[113] R. Z. Marandi, P. Madeleine, O. Omland, N. Vuillerme, and A. Samani, "Eye movement characteristics reflected fatigue development in both young and elderly individuals," *Sci. Rep.*, vol. 8, Sep. 2018, Art. no. 13148.

[114] J. Galbally, R. Haraksim, and L. Beslay, "A study of age and ageing in fingerprint biometrics," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 5, pp. 1351–1365, May 2019.

[115] D. Iakovakis, S. Hadjidimitriou, V. Charisis, S. Bostantzopoulou, Z. Katsarou, and L. J. Hadjileontiadis, "Touchscreen typing-pattern analysis for detecting fine motor skills decline in early-stage Parkinson's disease," *Sci. Rep.*, vol. 8, p. 7663, May 2018.

[116] X. Wang, T. Yu, M. Zeng, and P. Tague, "XREC: Behavior-based user recognition across mobile devices," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 1, no. 3, pp. 1–26, Sep. 2017.

[117] K. Cao, C. Wei, A. Gaidon, N. Arechiga, and T. Ma, "Learning imbalanced datasets with label-distribution-aware margin loss," in *Proc. 32nd Adv. Neural Inf. Process. Syst.*, 2019, pp. 1567–1578.

[118] S. Rahman, S. Khan, and F. Porikli, "A unified approach for conventional zero-shot, generalized zero-shot, and few-shot learning," *IEEE Trans. Image Process.*, vol. 27, no. 11, pp. 5652–5667, Nov. 2018.

[119] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[120] S. Al-Stouhi and C. K. Reddy, "Transfer learning for class imbalance problems with inadequate data," *Knowl. Inf. Syst.*, vol. 48, no. 1, pp. 201–228, Jul. 2016.

[121] C. X. Lu *et al.*, "Autonomous learning for face recognition in the wild via ambient wireless cues," in *Proc. World Wide Web Conf. (WWW)*, 2019, pp. 1175–1186.

[122] N. D. Lane, S. Bhattacharya, P. Georgiev, C. Forlivesi, and F. Kawsar, "An early resource characterization of deep learning on wearables, smartphones and Internet-of-Things devices," in *Proc. Int. Workshop Internet Things Towards Appl. (IoT-App)*, 2015, pp. 7–12.

[123] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," 2018. [Online]. Available: arXiv:abs/1803.01271.

[124] T. Yang, Y. Chen, and V. Sze, "Designing energy-efficient convolutional neural networks using energy-aware pruning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6071–6079.

[125] N. D. Lane *et al.*, "DeepX: A software accelerator for low-power deep learning inference on mobile devices," in *Proc. 15th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2016, pp. 1–12.

[126] S. Bhattacharya and N. D. Lane, "Sparsification and separation of deep learning layers for constrained resource inference on wearables," in *Proc. 14th ACM Conf. Embedded Netw. Sensor Syst. CD-ROM (SenSys)*, 2016, pp. 176–189.

[127] S. Han, J. Pool, J. Tran, and W. J. Dally, "Learning both weights and connections for efficient neural networks," in *Proc. 28th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, vol. 1, 2015, pp. 1135–1143.

[128] R. Appuswamy *et al.*, "Structured convolution matrices for energy-efficient deep learning," Jun. 2016. [Online]. Available: arXiv:1606.02407.

[129] Z. Lu, V. Sindhwani, and T. N. Sainath, "Learning compact recurrent neural networks," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 5960–5964.

[130] H. Pham, M. Y. Guan, B. Zoph, Q. V. Le, and J. Dean, "Efficient neural architecture search via parameter sharing," in *Proc. ICML*, 2018, pp. 4092–4101.

[131] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An extremely efficient convolutional neural network for mobile devices," in *Proc. Comput. Vis. Pattern Recognit.*, 2018, pp. 6848–6856.

[132] X. Dai *et al.*, "ChamNet: Towards efficient network design through platform-aware model adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 11398–11407.
[133] L. Duong, T. Cohn, S. Bird, and P. Cook, "Low resource dependency parsing: Cross-lingual parameter sharing in a neural network parser," in *Proc. 53rd Annu. Meeting Assoc. Comput. Linguist. 7th Int. Joint Conf. Nat. Lang. Process. Short Papers*, vol. 2, Jul. 2015, pp. 845–850.
[134] J. Chen and X. Ran, "Deep learning with edge computing: A review," *Proc. IEEE*, vol. 107, no. 8, pp. 1655–1674, Aug. 2019.
[135] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo, and J. Zhang, "Edge intelligence: Paving the last mile of artificial intelligence with edge computing," *Proc. IEEE*, vol. 107, no. 8, pp. 1738–1762, Aug. 2019.
[136] Y. Kang *et al.*, "Neurosurgeon: Collaborative intelligence between the cloud and mobile edge," *SIGARCH Comput. Archit. News*, vol. 45, no. 1, pp. 615–629, Apr. 2017.
[137] Z. Zhao, K. M. Barijough, and A. Gerstlauer, "DeepThings: Distributed adaptive deep learning inference on resource-constrained IoT edge clusters," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 11, pp. 2348–2359, Nov. 2018.
[138] M. Xu, F. Qian, M. Zhu, F. Huang, S. Pushp, and X. Liu, "DeepWear: Adaptive local offloading for on-wearable deep learning," *IEEE Trans. Mobile Comput.*, vol. 19, no. 2, pp. 314–330, Feb. 2020.
[139] T. Chen, S. Barbarossa, X. Wang, G. B. Giannakis, and Z. Zhang, "Learning and management for Internet of Things: Accounting for adaptivity and scalability," *Proc. IEEE*, vol. 107, no. 4, pp. 778–796, Apr. 2019.
[140] L. Zeng, E. Li, Z. Zhou, and X. Chen, "Boomerang: On-demand cooperative deep neural network inference for edge intelligence on the industrial Internet of Things," *IEEE Netw.*, vol. 33, no. 5, pp. 96–103, Sep. 2019.
[141] V. Nguyen, P. Lin, and R. Hwang, "Energy depletion attacks in low power wireless networks," *IEEE Access*, vol. 7, pp. 51915–51932, 2019.
[142] L. Kang and H. Shen, "Preventing battery attacks on electrical vehicles based on data-driven behavior modeling," in *Proc. 10th ACM/IEEE Int. Conf. Cyber Phys. Syst. (ICCPS)*, 2019, pp. 35–46.
[143] I. Tomić and J. A. McCann, "A survey of potential security issues in existing wireless sensor network protocols," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1910–1923, Dec. 2017.
[144] I. Farris, T. Taleb, Y. Khettab, and J. Song, "A survey on emerging SDN and NFV security mechanisms for IoT systems," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 812–837, 1st Quart., 2019.
[145] C. X. Lu *et al.*, "Snoopy: Sniffing your smartwatch passwords via deep sequence learning," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 1, no. 4, p. 152, Jan. 2018.
[146] H. Hu, C. Lin, C. Chang, and L. Chen, "Enhanced secure data backup scheme using multi-factor authentication," *IET Inf. Security*, vol. 13, no. 6, pp. 649–658, 2019.
[147] J. Zhao, X. Xie, X. Xu, and S. Sun, "Multi-view learning overview: Recent progress and new challenges," *Inf. Fusion*, vol. 38, pp. 43–54, Nov. 2017.
[148] Y. Li, M. Yang, and Z. Zhang, "A survey of multi-view representation learning," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 10, pp. 1863–1883, Oct. 2019.
[149] D. Kim, K. Chung, and K. Hong, "Person authentication using face, teeth and voice modalities for mobile device security," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, pp. 2678–2685, Nov. 2010.
[150] H. Crawford, K. Renaud, and T. Storer, "A framework for continuous, transparent mobile device authentication," *Comput. Security*, vol. 39, pp. 127–136, Nov. 2013.
[151] B. Zhou, J. Lohokare, R. Gao, and F. Ye, "EchoPrint: Two-factor authentication using acoustics and vision on smartphones," in *Proc. 24th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2018, pp. 321–336.
[152] H. Gomi, S. Yamaguchi, K. Tsubouchi, and N. Sasaya, "Towards authentication using multi-modal online activities," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. ACM Int. Symp. Wearable Comput. (UbiComp)*, 2017, pp. 37–40.
[153] P. Kumar, S. Mukherjee, R. Saini, P. Kaushik, P. P. Roy, and D. P. Dogra, "Multimodal gait recognition with inertial sensor data and video using evolutionary algorithm," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 5, pp. 956–965, May 2019.
[154] H. Feng, K. Fawaz, and K. G. Shin, "Continuous authentication for voice assistants," in *Proc. 23rd Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2017, pp. 343–355.
[155] S. Pradhan, W. Sun, G. Baig, and L. Qiu, "Combating replay attacks against voice assistants," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 3, no. 3, p. 100, Sep. 2019.
[156] J. Li *et al.*, "Robust face recognition with deep multi-view representation learning," in *Proc. 24th ACM Int. Conf. Multimedia (MM)*, 2016, pp. 1068–1072.
[157] C. L. P. Lim, W. L. Woo, S. S. Dlay, D. Wu, and B. Gao, "Deep multiview heartwave authentication," *IEEE Trans. Ind. Informat.*, vol. 15, no. 2, pp. 777–786, Feb. 2019.
[158] D. Hintze, E. Koch, S. Scholz, and R. Mayrhofer, "Location-based risk assessment for mobile authentication," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. Adjunct (UbiComp)*, 2016, pp. 85–88.
[159] A. Wójtowicz and K. Joachimiak, "Model for adaptable context-based biometric authentication for mobile devices," *Pers. Ubiquitous Comput.*, vol. 20, no. 2, pp. 195–207, Apr. 2016.
[160] D. Hintze *et al.*, "Cormorant: Ubiquitous risk-aware multi-modal biometric authentication across mobile devices," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 3, no. 3, pp. 1–23, Sep. 2019.
[161] E. Toch *et al.*, "The privacy implications of cyber security systems: A technological survey," *ACM Comput. Surveys*, vol. 51, no. 2, p. 36, Feb. 2018.
[162] Y. X. M. Tan, A. Iacovazzi, I. Homoliak, Y. Elovici, and A. Binder, "Adversarial attacks on remote user authentication using behavioral mouse dynamics," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2019, pp. 1–10.
[163] Q. He, Y. Xu, Z. Liu, J. He, Y. Sun, and R. Zhang, "A privacy-preserving Internet of Things device management scheme based on blockchain," *Int. J. Distrib. Sensor Netw.*, vol. 14, no. 11, pp. 1–12, 2018.
[164] S. Wang *et al.*, "Adaptive federated learning in resource constrained edge computing systems," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1205–1221, Jun. 2019.
[165] G. Xu, H. Li, S. Liu, K. Yang, and X. Lin, "VerifyNet: Secure and verifiable federated learning," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 911–926, 2020.
[166] H. Liu, H. Ning, Y. Zhang, and L. T. Yang, "Aggregated-proofs based privacy-preserving authentication for V2G networks in the smart grid," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1722–1733, Oct. 2012.
[167] H. Gao, Z. Ma, S. Luo, and Z. Wang, "BFR-MPC: A blockchain-based fair and robust multi-party computation scheme," *IEEE Access*, vol. 7, pp. 110439–110450, 2019.

Yunji Liang received the Ph.D. degree in computer science from Northwestern Polytechnical University, Xi'an, China, in 2016.

He is an Associate Professor with Northwestern Polytechnical University. From 2012 to 2017, he worked with the University of Arizona, Tucson, AZ, USA, as a Visiting Scholar and the Postdoctoral Researcher. His research interests include pervasive computing, social computing, and intelligent system.

Sagar Samtani received the Ph.D. degree in management information systems from the Artificial Intelligence Laboratory, University of Arizona, Tucson, AZ, USA, in 2018.

He is an Assistant Professor and the Grant Thornton Scholar with the Operations and Decision Technologies Department, Kelley School of Business, Indiana University, Tampa, FL, USA. His research interests include cyber threat intelligence, AI for cybersecurity, dark Web analytics, and interpretable deep learning.

Bin Guo received the Ph.D. degree in computer science from Keio University, Tokyo, Japan, in 2009.

He was a Postdoctoral Researcher with Institut TELECOM SudParis, Évry, France. He is a Professor with Northwestern Polytechnical University, Xi'an, China. His research interests include ubiquitous computing and mobile crowdsensing.

Zhiwen Yu (Senior Member, IEEE) received the Ph.D. degree in computer science from Northwestern Polytechnical University, Xi'an, China, in 2005.

He is a Professor with Northwestern Polytechnical University. He has worked as an Alexander Von Humboldt Fellow with Mannheim University, Mannheim, Germany, from November 2009 to October 2010, and as a Research Fellow with Kyoto University, Kyoto, Japan, from February 2007 to January 2009. His research interests cover ubiquitous computing and HCI.