Deep-Learning and Privacy Techniques for Data-Driven Soft Sensors

Subjects: Computer Science, Artificial Intelligence

Contributor: Razvan Bocu , Dorin Bocu , Maksim Iavich

The continuously increasing number of mobile devices actively being used in the world amounted to approximately 6.8 billion by 2022. Consequently, this implies a substantial increase in the amount of personal data collected, transported, processed, and stored. An integrated personal health data management system was designed and implemented, which considers data-driven software and hardware sensors, comprehensive data privacy techniques, and machine-learning-based algorithmic models.

data-driven soft sensors deep learning

1. Data Acquisition through Mobile Devices and Sensors

Mobile devices provide a comprehensive set of functional features, which may be used for the proper processing and collection of related data. As an example, modern smartphones are equipped with powerful hardware components, such as multicore processors, sophisticated mobile graphical processing units (GPU), several gigabytes of memory, and a comprehensive set of built-in sensors. Additionally, it is possible to add new sensors using the wireless and even wired connection features of these mobile devices. The following subsection presents relevant contributions, which pertain to the design and implementation of full privacy-preserving data channels. Moreover, the possibility to conduct arithmetic operations directly over the encrypted data is discussed.

1.1. Remarks Concerning Full Privacy-Preserving Data Computation

The authors of [1] reported a verifiable data processing model that is related to encrypted input data in connection with mHealth (mobile health) software systems. The algorithmic scheme that is designated as accumulation tree was reported in [2], which verifies the results of geographical proximity tests. Furthermore, ref. [3] described the results that relate to verifiable computation use cases, which pertain to encrypted input data. It is important to mention that most of the existing approaches consider data processing at the level of the client devices. This approach does not apply to integrated data management systems which consider personal private data.

The advantages of cloud-based data storage and processing are obvious [4]. However, the design of the proper data security approaches determines a significant problem that generates conceptual issues to the cloud service providers [5][6].

The implied service providers aim to design and deploy layered security mechanisms. Nevertheless, the plain text data may still be accessed and used through proper intrusion techniques. Consequently, data must be encrypted before transmission to the respective external data processing modules. The relevant reviewed papers suggest a significant computation overhead connected to the mobile client devices [7]. This is especially relevant for the personal mobile devices, which collect medical data processed by proper integrated software systems. There are, however, approaches [8] that do not specify proper data privacy mechanisms [9] when the data are transmitted through the respective data channels. The proper management of personal health information (PHI) data refers to ethical principles and formal regulations [10]. Thus, it is necessary to design and implement integrated data processing systems that consider all the relevant constraints. The authors of [11] described the general architectures and the life cycles of cloud-based data processing services.

Ever since C. Gentry first described the concept of homomorphic encryption in 2009 [12], significant research has focused on improving [13] this computationally expensive data processing scheme. Consequently, many relevant real-world use cases pertain to the use of proper powerful hardware resources [14]. Moreover, the initial homomorphic encryption approaches were particularly computationally expensive relative to the respective real-world use cases [15]. Furthermore, the algorithmic apparatus was improved through multiple development phases [16]. Some papers have reported improvements to the computational efficiency of homomorphic encryption. For example, refs. [17][18][19][20] expanded the initial set of algorithms. The algorithmic model presented in [21][22] and also in ref. [23] may be used during the design of data processing components that are part of relevant integrated data management systems [24]. Furthermore, it is important to note the papers [25][26] that are connected to the full scope of ubiquitous systems. Thus, ref. [27] described a software application defined by two functional requirements. First, the system is able to conduct the semantic analysis of data that are produced by user interactions, which are connected to various contextual parameters that determine usual activities of daily living (ADL). This has the goal of determining the relevant behavioral patterns that define complex activities. Moreover, the software system is based on an algorithmic routine that supports the decision-making processes. Furthermore, a relevant contribution is reported in papers [28][29]. Additionally, ref. [30] described a general architecture of a ubiquitous system that is compatible with general medical use case scenarios and data storage models, such as the ones that are described in papers [31][32]. Moreover, software systems defined by interesting architectural models are presented in papers [33][34], and also [35][36]. It is significant to note that the authors of papers [37][38], and also [39] propose technical solutions that are relevant for the implementation of distributed personal data processing systems, which use wireless data transfer channels. Furthermore, the survey effort that is included in [40] created interesting perspectives on related scientific problems. Moreover, the authors of [41][42] proposed interesting data transmission models relative to next-generation radio networks, while [43] described a versatile data communication channels management system, which can be used in a variety of real-world use case scenarios, including vehicular ad hoc networks (VANET).

Moreover, it is important to mention the contributions that were described in [44][45], considering that they presented one of the few existing integrated personal data management systems, which fully implements data protection mechanisms considering all the relevant stages: data collection, transportation, processing, and long-term storage.

The survey that was conducted suggests the following requirements for any suitable integrated personal data management system.

- The collection of personal data is conducted using mobile client devices.
- The data is transferred to central data processing components.
- The data are properly and securely stored, and privacy-preserving data is processed.
- The system should be specified considering a flexible and decoupled system architecture which would allow for an efficient extension and re-structuring of the system in the future.
- The legal and formal requirements that are formalized by American and European regulations are also considered.
- The efficient integration of the system in the target software frameworks considers the specifics of the respective use cases, as well as all the technical and legal requirements.

1.2. Analytical Remarks Concerning Similar Contributions

Thus, ref. [46] relates to a comprehensive review of similar data privacy mechanisms, with a focus on e-Health software systems. Relevant advantages and disadvantages of reviewed models are analyzed. The papers were selected considering the similarity that was observed in the reviewed literature. The authors also describe the general features of a technical standard, which may define an e-Health system. The paper also includes a taxonomy of cloud-based models, while the relevant personal data privacy and security requirements enforced by the Health Insurance Portability and Accountability Act (HIPAA) [47][48] are analyzed. It is important to note that the authors describe a secure and dependable system architecture, which is compatible with electronic health scenarios that could guarantee efficiency, reliability, and a properly regulated access framework to health information. The main drawback of this architecture is its inability to deploy on distributed and structurally scalable infrastructures. Additionally, only standard asymmetric encryption models are implemented, which do not provide the necessary degree of health data privacy.

The general scope of cloud-based healthcare computing has modified real-world healthcare in several ways. Cloud infrastructures provide a discernible advantage in the scalability of service, and the possibility to alter the related computational and data storage resources. Additionally, other articles examine the implied security and data privacy-preserving mechanisms. This is an important aspect of the overall research problem, as it determines important legal and technological aspects that should be evaluated. In this respect, ref. [49] examines several scientific approaches that miss at least some of the necessary technical features. Thus, it is important to mention the end-to-end private data transmission channels, the mandatory scalability, and the architectural compatibility of diverse technical platforms and frameworks, which concern the implied client and back-end (server) components.

The obvious advances in the field of information and communication technology naturally relate to an improved economic environment that offers higher-value services to consumers and businesses. The health sector benefits from this progress. Although the cloud-based system architectures provide clear advantages, the remaining security and data privacy issues should still be considered and addressed. Thus, ref. [50] presented a distributed system that considers various data security levels and data encryption models. This heterogeneous architectural structure implies administrative, functional, and data security problems that suggest that the reported approach is not suited for real-time deployments of large-scale medical data processing systems.

The continuous development of Internet of Things (IoT) as a theoretically and practically relevant paradigm, which has occurred during the past twenty years, implies that novel personal data management approaches may be developed. Thus, ref. [51] presented an important problem, which concerns the fully secure preservation of personal data privacy. The article proposes an access control mechanism for cloud-based data that follows a certificate-based authentication model. The authors describe the methodology of the approach using the results of experimental evaluation processes. This suggests an apparent enhancement of the overall system's security and performance through the optimization of the time needed to specify and implement the data and service access permissions. Nevertheless, the proposed approach does not offer the necessary scalability or end-to-end private data transmission channels between the client devices and the back end data processing components.

Significant progress has been made in the scope of cloud-based healthcare applications in the past ten years, particularly due to the implied remote access features, among other advantages. It is important to note that the reviewed literature demonstrates the resistance of certain end users to the adoption of the new technologies, particularly in developing nations [52].

Attribute-based encryption (ABE) models represent an interesting use case in healthcare. Patients encrypt their electronic health record (EHR), assign the attributes, and send them to the cloud. Healthcare professionals receive the encrypted EHR corresponding to their field of expertise from the cloud-based system. Decryption of the EHR data presumes that the medical personnel receive the secret keys from the key generation center (KGC). Thus, the KGC stores the secret keys of all the encrypted EHR records. Consequently, it is possible to decrypt the relevant patients' records, which represents a security issue. A decentralized ABE scheme addresses this issue, but it implies significant computation and communication costs. Furthermore, unauthorized medical employees may be able to read the patients' private EHR data. Additionally, the privacy of the KGC's secret keys and the doctor's attribute privacy determine relevant research aspects. Thus, ref. [53] presented a cloud-based privacy-preserving e-health (CP2EH) scheme, which addresses the issues of unauthorized access to patient records and the proper management of the doctor's attribute privacy relative to an ABE scheme. The presented model includes the oblivious transfer (OT) and zero-knowledge proof (ZKP) protocols in the centralized ABE scheme. Thus, the OT protocol ensures the privacy of the secret keys and the doctor's attribute. Despite the reported advantages, the system is selective concerning the accepted data acquisition devices. Moreover, it is compatible with only certain software frameworks, it does not scale well, and it does not implement end-to-end secure data transmission channels.

The authors in [54] presented an attribute-based encryption (ABE) access control model. This enforces controlled and possibly multi-level access delegation policies. Moreover, the authors evaluate the possibility of deploying such a system in an e-health environment with the goal of safely sharing EHR data of the patients enrolled in the system. The authors assert that the proposed mechanism is safe from some plaintext attacks and from attacks based on attribute collusion [55].

2. General Mobile Collection of Sensitive Data

Mobile devices possess the hardware capabilities needed to facilitate general data collection and processing. These include powerful multicore central processing units (CPU), graphical processing units (GPU), and random access memory (RAM) which sustain powerful and versatile operating systems. The mentioned hardware and software features support efficient data sensing and collection operations, together with the usual smartphone core functions.

Consequently, the built-in mobile sensors are able to collect data considering an adequate frequency for the data acquisition interval and for the private data categories.

The remarks are applicable to many types of mobile wearable devices, such as smartwatches, which can be assimilated to the wider scope of Internet of Things (IoT) devices, as long as they are connected to the Internet or are linked to devices that are directly connected to the Internet [56]. These devices are rapidly becoming capable of performing complex measurements, and even local data analysis processes [57]. In principle, mobile device manufacturers implement and provide the required mobile applications, which can be installed on their wearable devices. Nevertheless, although these mobile applications are adequate for general use case scenarios, specialized applications are required to sustain specific real-world scenarios.

Motion sensors are designed to measure both the rotational and acceleration forces over the three axes of the related device. Thus, the hardware motion sensors keep track of the angular velocity and acceleration, and the software sensors may produce an output according to either a continuous or an event-driven pattern. Moreover, the position sensors imply the measurement of changes in the Earth's magnetic field related to the actual physical orientation, while environmental sensors are typically activated by an event and return a value measurement in the form of one scalar. These sensors may be configured to return continuous measurements, which may attain a frequency of approximately 200 Hz, while their power consumption is still kept at a low level [46].

Certain measurements that discern biological and physiological parameters are supported on particular mobile devices due to specialized health sensors. As an example, many mobile devices, including smartwatches, feature optical sensors used to detect the changes in the volume of the blood flowing through the arteries. Consequently, physiological heart parameters are evaluated. Additionally, studies that pertain to other health problems, such as sleep disorders in the scope of polysomnography, also use sensors [58][59].

Data generated by user interaction with a touchscreen can be quantified by the number of “keystrokes” [60] or by analyzing the touch data generated by the user [61]. Considering the former situation, the virtual keystrokes are recorded, and the timestamp and pressure data are also logged for each keystroke. The acquired data allow analysts to discern even more complex features, such as the time between keystrokes, the time allocated for touch and hold operations, and so on [62]. Supplementary to the actual keystrokes, modern touchscreen panels significantly expanded the user interaction area, which includes the screen zones that are sensitive to user touch operations. Thus, it is possible to precisely determine the location of the touch points using a coordinate system relative to the X and Y coordinates of the screen. Additionally, all of the other usual gestures, such as pinch, tap, swipe, multitouch, and more advanced user interaction parameters, such as angle, velocity, trajectory, and acceleration, can be extracted [63].

Data connections represent a basic but fundamental aspect of mobile devices that imply the implementation and full or partial compatibility with a vast set of network protocols. The networked data connections generate private data patterns concerning the user’s daily patterns. Consequently, they can support the profiling of human behaviour and the acquisition of related sensitive personal data. Considering that the 5G radio standard is currently during its early stages of commercial deployment and that the 6G radio standard is under development, it can be asserted that the improved data transfer rates and the significantly lower latency values will expand the functional capabilities of machine-to-machine (M2M) communications. This should essentially increase the research and commercial relevance of mobile devices [64].

3. Real-World Sensors Use Case Scenarios

The two mainstream mobile operating systems, Android and iOS, initially offered less than 500 applications for download in their application stores. Currently, Google Play, which represents the Android applications store, includes over 3.5 million applications, while Apple’s App Store offers approximately 2.2 million applications. It is also interesting to note that Amazon App store contains approximately 500,000 applications [65]. The extensive range of applications cover various use case scenarios, the most relevant of which are discussed in the following paragraphs.

3.1. User Authentication Systems

Considering the mainstream user authentication systems, legitimate users are required to provide a secret token, such as a password or a personal identification number (PIN) code. This authentication model is commonly known as “what you know”. Moreover, there are authentication systems based on certain physical items, such as public key infrastructure (PKI) cards, which are known as “what you have”. Additionally, other authentication systems consider users’ physical features, such as fingerprints or geometry of the eyes, to perform the authentication. This is known as the “what you are” paradigm [66].

Biometrics are common and fundamental instruments of mobile authentication systems. The biometrics may belong to both physiological and behavioral categories [67]. Nevertheless, using such authentication models implies

that the device remains unlocked and accessible, and any unauthorized access is possible. This shortcoming may be averted through continuous authentication schemes relative to mobile devices, which use behavioral biometric authentication mechanisms [68]. Thus, the biometric data are continuously collected through a passive model during normal use of the mobile device, which ensures that the user's physical features correspond to those of the legitimate owner. Nevertheless, several logical or environmental features, such as scenarios, modalities, or environmental traits, may adversely influence the accuracy of mobile biometric systems [69]. Thus, the literature reports hybrid solutions, which combine background sensors [70][71], touchscreen devices [72], and network information [73]. This supports the development of higher-accuracy continuous authentication systems, which are based on behavioural biometric mechanisms.

3.2. Fitness and Healthcare Systems and Services

Mobile applications and devices play an important role in the healthcare sector. Thus, "mHealth" (mobile health) is a concept referring to a subset of eHealth that encompasses medical and public health procedures supported by mobile devices. Mobile applications support the general healthcare processes. Thus, patients may avail themselves of improved and more efficient services regarding acute and chronic conditions [74].

Mobile applications can represent real-world use cases, which are capable of analyzing body postures and generating reports concerning mental disorders [75]. They may also monitor medical conditions, such as Parkinson disease, stress, dementia, among others [76][77]. Furthermore, mobile health applications may support the improvement of a healthy lifestyle. Thus, a variety of mobile devices, such as mobile phones and smartwatches, are used to track the intensity of the measured physical activity, including all the relevant physiological parameters [78][79][80].

Additionally, existing scientific studies report integrated mHealth and eHealth software systems that support the collection of personal health data using mobile and wearable devices, the processing of the data components, the format of the encrypted data to conduct arithmetic operations, and secure-long term personal health data storage. This type of full privacy preserving approach, which relates to homomorphic encryption and virtualized 5G data channels, was described in [44][45].

3.3. Services Based on Location Data

Mobile devices fetch geolocation data using several sources, including the Global Positioning System (GPS) hardware devices. These data are used by mobile applications to determine the geographical position of the users for a variety of purposes, such as navigation hints data or targeted advertising [81]. The applications that consider geolocation data (location-aware applications) belong to the realm of the context awareness paradigm [82]. Moreover, radio protocols used to transmit data short distances, such as Wi-Fi (Wireless Fidelity) and Bluetooth, allow the mobile devices to exchange data with neighbouring devices and consequently use them for their purposes. This approach may be used to specify a semantic context, which is determined by the immediate environment. As an example, the contribution that is reported in [83] described the specification and implementation

of virtual tours in museums, which would provide relevant information to the visitors based on the neighbourhood of the visitors' actual position in the museum. Furthermore, interesting relevant aspects may also be studied in [84].

3.4. Remarks Concerning Other Relevant Use Cases

Considering the mainstream use cases, background sensors improve the end users' experience in various ways. As an example, the determination of a mobile device's position is facilitated by the background sensors, which implement the automatic change of the screen orientation. Obviously, data generated by the light sensors support the automatic adjustment of the screen brightness. Moreover, the proximity sensor manages the screen lock or unlock states in different situations, for example when placing a phone call. Another interesting use case is represented by the augmented reality (AR) applications in fields such as entertainment, commerce, and navigation [85]. The AR applications rely essentially on the data generated by the background sensors.

The ubiquity of modern mobile devices allows for more complex but useful real-world use case scenarios, such as mobile participatory sensing [86]. Thus, particular users voluntarily agree to share their devices to collect data that are relevant for the analysis of various aspects of the implied reality. This mediates the collection of relevant data, which are consequently used to assess, measure, and map various phenomena through a crowd-sourced participatory manner [81]. These use case scenarios include, among others, monitoring urban noise and pollution levels, monitoring urban cleanliness levels, and monitoring urban road and traffic conditions [87].

4. Proper Management of Sensitive Private Data

The automated management of data collected through mobile device use involves interaction with an appreciable amount of sensitive private data. It is important to note that some mobile sensors, such as GPS hardware components, microphones, and cameras, are especially difficult to tamper with, as they require special access permissions. Nevertheless, other mobile sensors, devices, or resources, such as the touchscreen, accelerometer, and networking data logs, require a lower level of access permission. Additionally, these data may be used to create a backdoor to sensitive personal data, considering that they can be sufficient to re-identify a particular individual through attributes, such as personal health data, particulars of daily routines, or demographic data.

The intimate nature of sensitive personal data requires the design and implementation of particular secure data management mechanisms. The most defining trait of this type of data relates to its uniqueness relative to the respective individuals. This is particularly relevant in relation to biometric data. Considering the wider scope of biometrics research, the main research and development challenges are represented by the mechanisms for storing personal data, the administrator or owner of the implied software and hardware data processing system, and the biometric features used to perform the authentication. Furthermore, the type and time reliability of the considered biometric features also represent a relevant question [88]. The next subsections discuss on the most relevant types of sensitive personal data that can be generated by the mobile devices' sensors.

4.1. Demographic Data

4.1.1. Sensors That Detect Movement

The authors in [89] considered the determination of a user's age range using data generated by an accelerometer. This was achieved during an experiment that involved performing a preset series of taps on a touchscreen relative to several contact spots. The experiment used the k-nearest neighbor (k-NN) algorithm, which produces an accuracy of 85.3%. Moreover, the authors of [90] reported an algorithmic mode that discriminates an adult from a child through behavioural particularities captured by the mobile motion sensors. The main hypothesis states that children, who have smaller hands, are shakier. The algorithmic model produced an accuracy of 96% through the random forest (RF) approach. The scientific contribution reported in [91], obtained the gender of the end users by analyzing their their walking routines data, which were collected by mobile motion sensors. The proposed model produced an accuracy of 76.8% using support vector machines (SVMs), and bagging algorithms. Moreover, the authors of [92] described an approach for recognition of gender data using gait (walking) data, which were collected by the mobile sensors. The reported accuracy was 96.3%, and the process used the bagged tree classifier.

The authors of [93] reported an automatic gender recognition algorithm, which uses the data collected by a gyroscope and accelerometer. The generated accuracy was 80% using the principal component analysis (PCA) technique. Moreover, the authors of [94] determined gender and age data using hidden Markov models (HMMs). Thus, the authors set up a competition which compared data collected by an accelerometer with gyroscope data using the respective mobile devices. The reported error percentage was 24.23% relative to the gender and 5.39% relative to age. The notable progress in the field of deep learning enhanced the results, as was the case with the findings described in [95]. Thus, the authors mentioned an accuracy of 94.11%, which was obtained through the analysis of gait (walking) data as it related to gender classification. The authors used long short-term memory (LSTM) and recurrent neural networks (RNNs) which are suitable for capturing the temporal dependencies that defined by the analyzed data.

4.1.2. Touchscreen Data

In [96], the authors categorized end users in two categories, adults and children, based on the mechanics of tap and swipe gestures. The authors describes an active user detection (AUD) algorithm, which generates an accuracy of 97%. Furthermore, ref. [97] presented a database that stores childrens' mobile interaction data. The considered touch interaction data allowed the children to be assigned to three categories, which included ages from 18 months to 8 years. The described model was based on the support vector machine (SVM) technique and yielded an accuracy of 90.45%. Furthermore, the authors of [90] reported a study based on the random forest (RF) technique, which used the tap gesture data to distinguish between adults and children. The model functions with an accuracy of 99%. Other papers report on using touchscreen data to determine an individual's gender. The study reported in [98] considered the prediction of soft biometrics data generated by swipe gestures. The measured accuracy was 78%, which was based on a decision voting scheme determined by four distinct classifiers: decision tree (DT), naive Bayes (NB), support vector machine (SVM), and logistic regression (LR). The authors of [99] collected behavioral data using mobile devices' accelerometers, gyroscopes, and orientation sensors, which were activated

during the end users' interactions with their mobile devices. The gesture data, which determine the gender of the user, were processed using a k-NN classifier with an accuracy of 93.65%.

4.1.3. Sensor Data Related to Mobile Applications, Location, and Network

Research has proven a correlation between geolocation data and the end users' demographics and usage patterns. As an example, in [100], the researchers stressed the significance of data generated by mobile devices in the context of demographic modeling and data measurement, while circumventing the need for traditional censuses and sociological research. This approach significantly speeds up the related political decisions. Furthermore, the authors of [101] considered radius, eccentricity, and entropy as three parameters that define travel behavior. More precisely, the authors attempted to explain the correlation between mobile device use and personal travel behaviour, which further analyzes the correlation between the frequency of the phone calls, and certain demographic factors, such as age, gender, and the defining features of the environment.

Moreover, ref. [102] described an unsupervised, data-driven model designed to create user categories that consider high-resolution mobility data, which are acquired through mobile navigation applications. The contribution reported in [103] described a method for the inference of demographic information using social networks photos, which include geographic tagging data. More precisely, this shows how an individual's ethnic characteristics can be obtained from collected geolocation data related to two particular metropolitan zones. The described model determines three ethnic groups, and the accuracy was reported as 72% using logistic regression (LR).

The scientific contribution reported in [104] discussed the suitability of geolocation data in inferring information regarding marital status and actual residence. The described research process considered the determination of spatial and temporal features using human mobility patterns, together with other features related to the geographical context. This approach offers information concerning the places visited by the individuals under analysis, such as private homes, hospitals, or leisure facilities. The obtained accuracy was 80% based on an eXtreme gradient boosting (XGBoost) algorithm [105]. The scientific presentation in [106] started with an analysis of gender-related behavioral patterns determined by mobile applications, which are related to the use of Wi-Fi and Bluetooth. The authors reported on the possibility to predict the gender of the end users and showed an accuracy of 91.8%. The algorithm used random forest (RF) and multinomial naive Bayes (NB). The data were collected from network connection logs, and the events were sorted according to occurrence frequency. An assessment of the temporal patterns was conducted relative to the 1000 events that occurred with the highest frequency. This type of contextual behavioral information is particularly useful in various domains, such as advertisement customization and the personalization of home screens.

4.2. Remarks Concerning the Study of Human Behaviour

The literature proves that the general patterns of users' daily activities and behavioural traits can be inferred from the data collected by mobile sensors [107]. This generates obvious problems regarding the privacy of the collected personal data, which should be properly addressed by academic and industrial research projects.

4.2.1. Motion Sensors

The authors of [108] described a system that is able to assess an individual's spatial mobility status. Thus, it can evaluate whether the person is stationary, walking, running, riding a bicycle, climbing stairs, going downstairs, or driving using only the accelerometer information. Their algorithmic approach, which is based on a support vector machine (SVM) technique, functions with an accuracy of up to 93.2%. Furthermore, the authors of [78] used mobile gyroscope and accelerometer data and developed an application used to track the user's daily routines. Their model is based on a decision tree (DT) classifier, and the average area under the receiver operating characteristic (AUROC) curve was over 99.0%.

The authors in [109] considered users' mobility while they were eating, as they are detected by the accelerometer sensor installed on smartwatches. The authors of [110] performed a classification of human drinking behavior. This took into account the data acquired by the accelerometer sensors of the mobile phones young adults used during nightlife activities. The accuracy of 76.1% was based on a density-based spatial clustering of applications (DBSCAN) algorithm. The respective approach also assessed the amount of ingested alcohol.

The assessment of user mood and physical state (sober, tipsy, or drunk) was conducted using the approach reported in [111] using accelerometer data. It also included a channel for users to report their own behaviour. Naturally, this was an auxiliary feature, which may not be regarded as an objective source of data. The algorithmic core was based on the random forest (RF) model, with an accuracy of 70%. Furthermore, mobile motion sensors were also used to collect data related to sleep, such as sleep habits and postures. The contribution that was reported in [112] uses accelerometer, gyroscope, and orientation data, which are retrieved using a smartwatch to detect and assess sleep postures (supine, left lateral, right lateral, prone). The reported algorithmic model produced an accuracy beyond 95%, which considered Euclidean distances. The described approach also evaluated the position of the users' hand considering the following three states: placed on the abdomen, chest, or head. The described model used a k-NN algorithm, with an accuracy greater than 88%.

4.2.2. Sensor Data Related to Mobile Applications, Location, and Network

The authors of [113] used GPS data to assess whether the user was standing, walking, or using other means of transportation. The algorithm used a fuzzy classifier, which calculated the speed and angle of the person relative to the ground. The measured accuracy was 96% considering the data, which were collected at five-second intervals. Additionally, it is also important to note that radio receivers and transmitters, by their nature, are also capable of providing information about users' behavioural patterns. This is also susceptible to generating sensible personal data security issues, which should be addressed. Thus, ref. [114] used the received signal strength indicator (RSSI) to determine user activity types. These were selected from the following set of states: lying down, falling, walking, running, sitting down, and standing up. The algorithmic model used a convolutional neural network (CNN), and the accuracy rate was 97.7%. The authors of [115] used three neural networks relative to the channel state information (CSI), which was measured by the Wi-Fi module. This technique can allegedly determine whether an individual is sitting, standing, or walking with an accuracy rate of 83%.

4.3. Remarks Regarding Body Features and Health Parameters

4.3.1. Motion Sensors

The body mass index (BMI) is a mathematical ratio that correlates the body mass and height of any person. The classic modality to compute this index is providing weight and height using the formula to calculate BMI. Human gait or style of walking is sustained by the synergistic cooperation established between hundreds of muscles and joints. Consequently, mobile motion sensors are capable of discerning various muscle movements, which are transformed into specific patterns for the traits of the individuals, such as BMI. Thus, the authors of [116] proposed a hybrid model based on a convolutional neural network and long short-term memory (CNN-LSTM) architecture. This is able to estimate BMI using the data generated by the accelerometer and the gyroscope, and the maximum determined accuracy is 94.8%. Considering BMI as a reference, several other health attributes may be determined [117][118]. It is interesting to note that another physiological variable that can be evaluated using accelerometer data is the level of stress. Thus, the authors of [119] reported an accuracy of 71% using the mentioned techniques and also the naive Bayes algorithm.

4.3.2. Remarks Concerning the Touchscreen

Data generated by mobile sensors may be used to assess, even to diagnose, certain medical conditions. Thus, it is possible to determine whether a person suffers from Parkinson's disease through the analysis of the respective users' keystroke writing pattern, which is totally independent from the actual content of the text. The authors in [120] considered an SVM algorithm, which determines an area under the receiver operating characteristic (AUROC) of 0.88 relative to this particular problem. Furthermore, ref. [121] assessed several types of features, which are specified relative to various handwriting patterns. These are used as biometrics to study Parkinson's disease. Moreover, in [122], the authors demonstrated that people with longer thumbs require less time to conduct swipe gestures.

4.3.3. Sensors Data Related to Mobile Applications, Location, and Network

The authors of [123] described an application that detects periods of psychological depression using geolocation patterns, which are retrieved from the mobile devices of individuals with bipolar disorder (BD). The model uses a linear regression algorithm, together with a quadratic discriminant analysis algorithm. The method produced an accuracy of 85%. GPS data may also be used to detect various sleep disorders, such as sleep–wake stages and sleep-disordered breathing disorders (SRBD), such as obstructive sleep apnea (OSA). The model uses SVM algorithms and demonstrated accuracy of up to 92.3% [124][125]. StayActive3 is an application that detects stress by analyzing the behavior of users via smartphone, using the data from the Wi-Fi, step counter, location, and battery level, among others. It is also worth mentioning the software system, which is referred to as StayActive [126]. The authors used a combination of simple relaxation scores that relate to the information acquired from the sleeping patterns of enrolled users. This analysis measures the longest time intervals during which the enrolled end users did not touch the screen, the patterns of their social interaction, and physical activity to evaluate the level of the stress.

4.4. The Detection of Psychological Mood and Emotions

End users' daily activities are dependent on their psychological mood. Consequently, valuable related data may be collected by various sensors.

4.4.1. Motion Sensors

The authors of [127] researched the influence that mood may have on the recognition accuracy rate of mobile biometric systems. Thus, by using an RF classifier, the authors discovered users with face recognition accuracy less than 70% exhibited the fewest psychological mood changes. The accelerometer provides useful data concerning users' walk patterns, which can be used to assess psychological mood relative to the following three states: happy, sad, or neutral. It is worth noting that the authors of [128] assessed mood using an RF algorithm, which produced a mean AUROC of 81%.

4.4.2. Touchscreen Data

Many studies demonstrate a correlation between users' interaction patterns with the screens of their mobile devices and their psychological mood. Thus, the authors of [129] researched the development of psychiatric diseases using an unobtrusive setup deployed in the patients' personal environment. The process explored the connection between bipolar affective disorder syndrome and the use of mobile devices. Considering the data generated by keystroke metadata and the accelerometer sensor, they obtained a detection accuracy of 90.31% relative to the proper detection of psychiatric conditions. The findings reported in [130] described a preventive medical treatment recommendation system, which may be useful to prevent the actual onset of clinical depression. Thus, the authors presented a mobile application, which was used to acquire the users' psychological states through the analysis of data provided by the call logs and the applications' usage history. The model produced an accuracy score of 86%.

The analysis of finger strokes patterns during games [131] can help distinguish between four emotional states: excited, relaxed, frustrated, and bored. The SVM algorithm produced an accuracy score of 69%. Moreover, the findings reported in [132] analyzed the pattern of finger strokes as an indication of the end user's psychological state, which can be classified as on e of three possible values: positive, negative, or neutral. The detection performed with an accuracy of 90.47% relative to a linear regression model.

4.4.3. Sensors Data Related to Mobile Applications, Location, and Network

MoodExplorer is an application that collects data using various mobile sensors, such as GPS, accelerometer, and Wi-Fi components [133]. The authors inferred the correlation established between psychological states, which were reported by the end users themselves, and the usage patterns of the respective mobile devices. The reported approach determines five types of emotions: happiness, sadness, anger, surprise, and fear. The algorithmic model is called Graph Factor. The performance was evaluated using a metric designated as "match", which featured an average value of 62.9%.

4.5. User Tracking through Location Data

Although mobile devices often feature dedicated GPS location devices, it is possible to determine geographic location using the data generated by other mobile sensors.

4.5.1. Motion Sensors

Certain scientific articles demonstrate that the geographic location of a person can be determined using data generated by several mobile sensors, such as accelerometer, gyroscope, and magnetometer, during the person's daily routines that involve using public transport, walking, or driving. The authors of [134] comparatively analyzed pre-defined routes, which were used by the end users relative to different means of transportation, such as walking, train, bus, or taxi. They compared the routes using a dynamic time warping (DTW) algorithm, which generated a Kullback–Leibler distance of 0.00057 relative to a taxi trip.

The authors in [135] described a modality that uses accelerometer data to track the end users' underground routes. The generated accuracy was 92% considering six visited underground stations, which were based on boosted naive Bayes (NB), and decision tree (DT) algorithms. The authors of [136] proposed an algorithmic model that determines the geographic location of vehicle drivers using the data generated by mobile motion sensors. The described approach considers an approximation of the related trajectory using accelerometer data. The map coordinates are correlated with the approximated trajectory to generate precise geographic location data. The approach that is presented allows for the end user to be located with a maximum error of 200 m. The distance is calculated as the radius between the center of the circle, which represents the actual person's location, and the approximated geographical location.

4.5.2. Sensor Data Related to Mobile Applications, Location, and Network

The end users' geographic location may also be determined using the data that identify encountered Wi-Fi networks. Thus, the authors of [137] described the indoor determination of the end users' location in a real-time fashion. The geographical location determination was conducted with an accuracy of 85.7% through the utilization of a random forest (RF) algorithm.

4.6. Logging Keystroke Data and Text Inference Using Motion Sensors

Touchlogger [138] is an application that aims to detect the precise zone of the screen that is touched. The process considers the device's micromovements as they are detected by the mobile gyroscope and accelerometer. The proposed approach considers a division of the screen into ten zones, which are analyzed using a probability density function relative to a Gaussian distribution. The application has shown an accuracy of 70%. It is also possible to determine the text that the end user generates based on the screen zones that are touched.

Furthermore, ref. [139] a related system has an accuracy rate of 93%, by utilizing a hierarchical classification scheme. Additionally, ref. [140] describes a controlled environment, which is used to detect various text patterns that

are entered using the mobile devices' touchscreen. Thus, the PIN code was correctly identified in 43% of the cases, while the unlock pattern was correctly detected in 73% of the cases. The algorithmic core is based on a hybrid model, which considers logistic regression (LR), and hidden markov models (HMM).

References

1. Guo, L.; Fang, Y.; Li, M.; Li, P. Verifiable privacy-preserving monitoring for cloud-assisted mHealth systems. In Proceedings of the 2015 IEEE Conference on Computer Communications, Hong Kong, 26 April–1 May 2015; pp. 1026–1034.
2. Zhuo, G.; Jia, Q.; Guo, L.; Li, M.; Fang, Y. Privacy-preserving verifiable proximity test for location-based services. In Proceedings of the 2015 IEEE Global Communications Conference, San Diego, CA, USA, 6–10 December 2015; pp. 1–6.
3. Fiore, D.; Gennaro, R.; Pastro, V. Efficiently verifiable computation on encrypted data. In Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, Scottsdale, AZ, USA, 3–7 November 2014; pp. 844–855.
4. Rokade, A.; Singh, M.; Arora, S.K.; Nizeyimana, E. IOT-Based Medical Informatics Farming System with Predictive Data Analytics Using Supervised Machine Learning Algorithms. Comput. Math. Methods Med. 2022, 2022, 8434966.
5. Kuzu, M.; Saiful Islam, M.; Kantarcioglu, M. Efficient similarity search over encrypted data. In Proceedings of the 2012 IEEE International Conference on Data Engineering, Washington, DC, USA, 1–5 April 2012; pp. 1156–1167.
6. Kadu, A.; Singh, M.; Ogudo, K. A Novel Scheme for Classification of Epilepsy Using Machine Learning and a Fuzzy Inference System Based on Wearable-Sensor Health Parameters. Sustainability 2022, 14, 15079.
7. Cao, N.; Wang, C.; Li, M.; Ren, K.; Lou, W.; Kantarcioglu, M. Privacy-preserving multi-keyword ranked search over encrypted cloud data. IEEE Trans. Parallel Distrib. Syst. 2014, 25, 222–233.
8. Orencik, C.; Savas, E. An efficient privacy-preserving multi-keyword search over encrypted cloud data with ranking. J. Parallel Distrib. Databases 2014, 32, 119–160.
9. Yu, J.; Lu, P.; Zhu, Y.; Xue, G.; Li, M. Toward Secure Multikeyword Top-k Retrieval over Encrypted Cloud Data. IEEE Trans. Dependable Secur. Comput. 2013, 10, 239–250.
10. Boldyreva, A.; Chenette, N.; Lee, Y.; O'Neill, A. Order-preserving symmetric encryption. In Proceedings of the 28th Conference on Theory and Applications of Cryptography Techniques, Trondheim, Norway, 30 May–3 June 2009; pp. 224–241.

11. Breiter, G.; Behrendt, M. Life cycle and characteristics of services in the world of cloud computing. IBM J. Res. Dev. 2009, 53, 3:1–3:8.
12. Gentry, C. A Fully Homomorphic Encryption Scheme; Stanford University: Stanford, CA, USA, 2009.
13. Brakerski, Z.; Vaikuntanathan, V. Efficient fully homomorphic encryption from (standard) LWE. SIAM J. Comput. 2011, 43, 831–871.
14. van Dijk, M.; Gentry, C.; Halevi, S.; Vaikuntanathan, V. Fully homomorphic encryption over the integers. In Proceedings of the 2010 EUROCRYPT Conference, French Riviera, France, 30 May–3 June 2010; pp. 24–43.
15. Coron, J.; Mandal, A.; Naccache, D.; Tibouchi, M. Fully homomorphic encryption over the integers with shorter public keys. In Proceedings of the Annual Cryptology Conference, Santa Barbara, CA, USA, 14–18 August 2011; Springer: Berlin/Heidelberg, Germany, 2011; pp. 487–504.
16. Steffen, S.; Bichsel, B.; Baumgartner, R.; Vechev, M. ZeeStar: Private Smart Contracts by Homomorphic Encryption and Zero-knowledge Proofs. In Proceedings of the IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 23–25 May 2022.
17. Gentry, C.; Halevi, S.; Smart, N.P. Fully homomorphic encryption with polylog overhead. In Proceedings of the Annual Cryptology Conference, Santa Barbara, CA, USA, 14–18 August 2011; Springer: Berlin/Heidelberg, Germany, 2011; pp. 465–482.
18. Brakerski, Z.; Gentry, C.; Vaikuntanathan, V. Fully homomorphic encryption without bootstrapping. In Proceedings of the 3rd Innovations in Theoretical Computer Science Conference, Cambridge, MA, USA, 8–12 January 2012; pp. 309–325.
19. Gentry, C.; Sahai, A.; Waters, B. Homomorphic encryption from learning with errors: Conceptually-simpler, asymptotically-faster, attribute-based. In Proceedings of the Annual Cryptology Conference, Santa Barbara, CA, USA, 18–22 August 2013; Springer: Berlin/Heidelberg, Germany, 2013; pp. 75–92.
20. General Data Protection Regulation. 2022. Available online: https://gdprinfo.eu/ro (accessed on 29 November 2022).
21. Aljeraisy, A.; Barati, M.; Rana, O.; Perera, C. Privacy laws and privacy by design schemes for the Internet of Things: A developer’s perspective. ACM Comput. Surv. 2021, 54, 102.
22. Barth, S.; de Jong, M.D.T.; Junger, M.; Hartel, P.H.; Roppelt, J.C. Putting the privacy paradox to the test: Online privacy and security behaviors among users with technical knowledge, privacy awareness, and financial resources. Telemat. Inform. 2019, 41, 55–69.
23. European Commission. PriMa: Privacy Matters, H2020-MSCA-ITN-2019-860315. 2022. Available online: https://www.prima-itn.eu/ (accessed on 5 December 2022).

24. European Commission. TReSPAsS-ETN: TRaining in Secure and PrivAcy-Preserving Biometrics, H2020-MSCAITN-2019-860813. 2022. Available online: https://www.trespass-etn.eu/ (accessed on 4 November 2022).
25. Seo, H.J.; Kim, S.Y.; Sheen, S.S.; Cha, Y. e-Health Interventions for Community-Dwelling Type 2 Diabetes: A Scoping Review. Telemed. E-Health 2021, 27, 276–285.
26. El Benny, M.; Kabakian-Khasholian, T.; El-Jardali, F.; Bardus, M. Application of the eHealth literacy model in digital health interventions: Scoping review. J. Med. Internet Res. 2021, 23, e23473.
27. Thakur, N.; Han, C.Y. An Ambient Intelligence-Based Human Behavior Monitoring Framework for Ubiquitous Environments. Information 2021, 12, 81.
28. Suma, V. Wearable IoT based distributed framework for ubiquitous computing. J. Ubiquitous Comput. Commun. Technol. 2021, 3, 23–32.
29. IBM Cloud Infrastructure. 2022. Available online: https://www.ibm.com/cloud (accessed on 20 May 2022).
30. Mondragón Martínez, O.H.; Solarte Astaíza, Z.M. Architecture for the Creation of Ubiquitous Services Devoted to Health. Universidad Católica de Pereira. 2022. Available online: http://hdl.handle.net/10785/9861 (accessed on 10 May 2022).
31. IBM Cloudant Storage Service. 2022. Available online: https://www.ibm.com/cloud/cloudant (accessed on 22 May 2022).
32. Apache OpenWhisk Service. 2022. Available online: ttps://developer.ibm.com/components/apache-openwhisk (accessed on 30 May 2022).
33. Akyildiz, I.F.; Wang, P.; Lin, S.C. SoftAir: A software defined networking architecture for 5G wireless systems. Comput. Netw. 2015, 85, 1–18.
34. Xia, X.; Xu, K.; Wang, Y.; Xu, Y. A 5G-Enabling Technology: Benefits, Feasibility, and Limitations of In-Band Full-Duplex mMIMO. IEEE Veh. Technol. Mag. 2018, 13, 81–90.
35. Boulogeorgos, A.-A.A.; Alexiou, A.; Merkle, T.; Schubert, C.; Elschner, R.; Katsiotis, A.; Stavrianos, P.; Kritharidis, D.; Chartsias, P.-K.; Kokkoniemi, J.; et al. Terahertz Technologies to Deliver Optical Network Quality of Experience in Wireless Systems Beyond 5G. IEEE Commun. Mag. 2018, 56, 144–151.
36. Kal, B.; Hamdaoui, B.; Guizani, M. Extracting and Exploiting Inherent Sparsity for Efficient IoT Support in 5G: Challenges and Potential Solutions. IEEE Wirel. Commun. 2017, 24, 68–73.
37. Simsek, M.; Aijaz, A.; Dohler, M.; Sachs, J.; Fettweis, G. 5G-Enabled Tactile Internet. IEEE J. Sel. Areas Commun. 2016, 34, 460–473.

38. Xu, L.; Collier, R.; O'Hare, G.M.P. A Survey of Clustering Techniques in WSNs and Consideration of the Challenges of Applying Such to 5G IoT Scenarios. IEEE Internet Things J. 2017, 4, 1229–1249.
39. Sekander, S.; Tabassum, H.; Hossain, E. Multi-Tier Drone Architecture for 5G/B5G Cellular Networks: Challenges, Trends, and Prospects. IEEE Commun. Mag. 2018, 56, 96–103.
40. Dhyani, K.; Bhachawat, S.; Prabhu, J.; Kumar, M.S. A Novel Survey on Ubiquitous Computing. In Data Intelligence and Cognitive Informatics; Springer: Singapore, 2022; pp. 109–123.
41. Hassan, M.; Singh, M.; Hamid, K.; Saeed, R.; Abdelhaq, M.; Alsaqour, R. Design of Power Location Coefficient System for 6G Downlink Cooperative NOMA Network. Energies 2022, 15, 6996.
42. Bolla, S.; Singh, M. Energy Harvesting Technique for Massive MIMO Wireless Communication Networks. J. Phys. Conf. Ser. 2022, 2327, 012059.
43. Marwah, G.P.K.; Jain, A.; Malik, P.K.; Singh, M.; Tanwar, S.; Safirescu, C.O.; Mihaltan, T.C.; Sharma, R.; Alkhayyat, A. An Improved Machine Learning Model with Hybrid Technique in VANET for Robust Communication. Mathematics 2022, 10, 4030.
44. Bocu, R.; Costache, C. A homomorphic encryption-based system for securely managing personal health metrics data. IBM J. Res. Dev. 2018, 62, 1:1–1:10.
45. Bocu, R.; Vasilescu, A.; Duca Iliescu, D.M. Personal Health Metrics Data Management Using Symmetric 5G Data Channels. Symmetry 2022, 14, 1387.
46. Acien, A.; Morales, A.; Fierrez, J.; Vera-Rodriguez, R.; Delgado-Mohatar, O. Becaptcha: Bot detection in smartphone interaction using touchscreen biometrics and mobile sensors. arXiv 2020, arXiv:2005.13655.
47. Hsieh, Y.P.; Lee, K.C.; Lee, T.F.; Su, G.J. Extended Chaotic-Map-Based User Authentication and Key Agreement for HIPAA Privacy/Security Regulations. Appl. Sci. 2022, 12, 5701.
48. Cohen, I.G.; Mello, M.M. HIPAA and protecting health information in the 21st century. JAMA 2018, 320, 231–232.
49. Sivan, R.; Zukarnain, Z.A. Security and Privacy in Cloud-Based E-Health System. Symmetry 2021, 13, 742.
50. Madan, S. Privacy-Preserved Access Control in E-Health Cloud-Based System. In Disruptive Technologies for Society 5.0; CRC Press: Boca Raton, FL, USA, 2021; pp. 145–162.
51. Daoud, W.B.; Meddeb-Makhlouf, A.; Zarai, F. A trust-based access control scheme for e-Health Cloud. In Proceedings of the 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA), Aqaba, Jordan, 28 October–1 November 2018; pp. 1–7.

52. Idoga, P.E.; Toycan, M.; Nadiri, H.; Çelebi, E. Factors affecting the successful adoption of e-health cloud based health system from healthcare consumers' perspective. IEEE Access 2018, 6, 71216–71228.
53. Yadav, V.K.; Yadav, R.K.; Verma, S.; Venkatesan, S. CP2EH: A comprehensive privacy-preserving e-health scheme over cloud. J. Supercomput. 2022, 78, 2386–2416.
54. Pussewalage, H.S.G.; Oleshchuk, V. A Delegatable Attribute Based Encryption Scheme for a Collaborative E-health Cloud. IEEE Trans. Serv. Comput. 2022.
55. Rajkumar, N.; Kannan, E. Attribute-based collusion resistance in group-based cloud data sharing using LKH model. J. Circuits Syst. Comput. 2020, 29, 2030001.
56. Delgado-Mohatar, O.; Tolosana, R.; Fierrez, J.; Morales, A. Blockchain in the Internet of Things: Architectures and implementation. In Proceedings of the IEEE 44th Annual Computers, Software, and Applications Conference, Madrid, Spain, 13–17 July 2020; pp. 1072–1077.
57. John Dian, F.; Vahidnia, R.; Rahmati, A. Wearables and the Internet of Things (IoT), applications, opportunities, and challenges: A survey. IEEE Access 2020, 8, 69200–69211.
58. Chen, Z.; Lin, M.; Chen, F.; Lane, N.D.; Cardone, G.; Wang, R.; Li, T.; Chen, Y.; Choudhury, T.; Campbell, A.T. Unobtrusive sleep monitoring using smartphones. In Proceedings of the International Conference on Pervasive Computing Technologies for Healthcare and Workshops, Venice, Italy, 5–8 May 2013; pp. 145–152.
59. Tayfur, I.; Afacan, M.A. Reliability of smartphone measurements of vital parameters: A prospective study using a reference method. Am. J. Emerg. Med. 2019, 37, 1527–1530.
60. Morales, A.; Fierrez, J.; Tolosana, R.; Ortega-Garcia, J.; Galbally, J.; Gomez-Barrero, M.; Anjos, A.; Marcel, S. Keystroke biometrics ongoing competition. IEEE Access 2016, 4, 7736–7746.
61. Tolosana, R.; Vera-Rodriguez, R.; Fierrez, J.; Ortega-Garcia, J. BioTouchPass2: Touchscreen password biometrics using time-aligned recurrent neural networks. IEEE Trans. Inf. Forensics Secur. 2020, 15, 2616–2628.
62. Acien, A.; Morales, A.; Monaco, J.V.; Vera-Rodriguez, R.; Fierrez, J. TypeNet: Deep learning keystroke biometrics. arXiv 2021, arXiv:2101.05570.
63. Tramèr, F.; Boneh, D. BioTouchPass2: Differentially private learning needs better features (or much more data). arXiv 2020, arXiv:2011.11660.
64. David, K.; Berndt, H. 6G vision and requirements: Is there any need for beyond 5G? IEEE Veh. Technol. Mag. 2018, 13, 72–80.
65. Statista. Number of Apps Available in Leading App Stores as of 2nd Quarter 2022. 2022. Available online: https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/ (accessed on 4 November 2022).

66. O'Gorman, L. Comparing passwords, tokens, and biometrics for user authentication. Proc. IEEE 2003, 91, 2021–2040.
67. Jain, A.K.; Nandakumar, K.; Ross, A. 50 years of biometric research: Accomplishments, challenges, and opportunities. Pattern Recognit. Lett. 2016, 79, 80–105.
68. Patel, V.M.; Chellappa, R.; Chandra, D.; Barbello, B. Continuous user authentication on mobile devices: Recent progress and remaining challenges. IEEE Signal Process. Mag. 2016, 13, 49–61.
69. Boakes, M.; Guest, R.; Deravi, F.; Corsetti, B. Exploring mobile biometric performance through identification of core factors and relationships. IEEE Trans. Biom. Behav. Identity Sci. 2019, 1, 278–291.
70. Acien, A.; Morales, A.; Vera-Rodriguez, R.; Fierrez, J.; Tolosana, R. Multilock: Mobile active authentication based on multiple biometric and behavioral patterns. In Proceedings of the International Workshop on Multimodal Understanding and Learning for Embodied Applications, Nice, France, 15 October 2019.
71. Wan, C.; Wang, L.; Phoha, V.V. A survey on gait recognition. ACM Comput. Surv. 2018, 51, 89.
72. Santopietro, M.; Vera-Rodriguez, R.; Guest, R.; Morales, A.; Acien, A. Assessing the quality of swipe interactions for mobile biometric systems. In Proceedings of the IEEE International Joint Conference on Biometrics (IJCB'20), Houston, TX, USA, 28 September–1 October 2020; pp. 1–8.
73. Li, G.; Bours, P. Studying Wifi and accelerometer data based authentication method on mobile phones. In Proceedings of the International Conference on Biometric Engineering and Applications, Amsterdam, The Netherlands, 16–18 May 2018; pp. 18–23.
74. Nussbaum, R.; Kelly, C.; Quinby, E.; Mac, A.; Parmanto, B.; Dicianno, B.E. Systematic review of mobile health applications in rehabilitation. Arch. Phys. Med. Rehabil. 2019, 100, 115–127.
75. Gravenhorst, F.; Muaremi, A.; Bardram, J.; Grünerbl, A.; Mayora, O.; Wurzer, G.; Frost, M.; Osmani, V.; Arnrich, B.; Lukowicz, P.; et al. Mobile phones as medical devices in mental disorder treatment: An overview. Pers. Ubiquitous Comput. 2015, 19, 335–353.
76. Faundez-Zanuy, M.; Fierrez, J.; Ferrer, M.A.; Diaz, M.; Tolosana, R.; Plamondon, R. Handwriting biometrics: Applications and future trends in e-security and e-health. Cogn. Comput. 2020, 12, 940–953.
77. Majumder, S.; Deen, M.J. Smartphone sensors for health monitoring and diagnosis. Sensors 2019, 19, 2164.
78. Anjum, A.; Ilyas, M.U. Activity recognition using smartphone sensors. In Proceedings of the IEEE Consumer Communications and Networking Conference, Las Vegas, NV, USA, 11–14 January 2013; pp. 914–919.

79. Antar, A.D.; Ahmed, M.; Ahad, M. Challenges in sensor-based human activity recognition and a comparative analysis of benchmark datasets: A review. In Proceedings of the International Conference on Informatics, Electronics and Vision and International Conference on Imaging, Vision and Pattern Recognition (icIVPR'19), Washington, DC, USA, 26 April 2019; pp. 134–139.
80. Khan, S.; Parkinson, S.; Grant, L.; Liu, N.; Mcguire, S. Biometric systems utilising health data from wearable devices: Applications and future challenges in computer security. ACM Comput. Surv. 2020, 53, 85.
81. Haris, M.; Haddadi, H.; Hui, P. Privacy leakage in mobile computing: Tools, methods, and characteristics. arXiv 2014, arXiv:1410.4978.
82. Saha, D.; Mukherjee, A. Pervasive computing: A paradigm for the 21st century. Computer 2003, 36, 25–31.
83. Luca, D.G.; Alberto, M. From proximity to accurate indoor localization for context awareness in mobile museum guides. Int. J. Uncertainty Fuzziness Knowl. Based Syst. 2016, 20, 1002–1009.
84. De Capitani Di Vimercati, S.; Foresti, S.; Livraga, G.; Amarati, P. Data privacy: Definitions and techniques. Int. J. Uncertainty Fuzziness Knowl. Based Syst. 2012, 20, 793–817.
85. Kim, S.J.; Kang, S.; Choi, Y.; Choi, M.; Hong, M. Augmented-reality survey: From concept to application. KSII Trans. Internet Inf. Syst. 2017, 11, 982–1004.
86. Burke, J.; Estrin, D.; Hansen, M.; Parker, A.; Ramanathan, N.; Reddy, S.; Srivastava, M.B. Participatory Sensing; UCLA: Center for Embedded Network Sensing: Los Angeles, CA, USA, 2006.
87. Melo, G.; Oliveira, L.; Schneider, D.; de Souza, J. Towards an observatory for mobile participatory sensing applications. In Proceedings of the International Conference on Computer Supported Cooperative Work in Design, Wellington, New Zealand, 26–28 April 2017; pp. 305–312.
88. Labati, R.D.; Piuri, V.; Scotti, F. Biometric privacy protection: Guidelines and technologies. In Proceedings of the International Conference on E-Business and Telecommunications, Seville, Spain, 18–21 July 2011; pp. 3–19.
89. Davarci, E.; Soysal, B.; Erguler, I.; Aydin, S.O.; Dincer, O.; Anarim, E. Age group detection using smartphone motion sensors. In Proceedings of the European Signal Processing Conference, Kos, Greece, 28 August 28–2 September 2017.
90. Nguyen, T.; Roy, A.; Memon, N. Kid on the phone! Toward automatic detection of children on mobile devices. Comput. Secur. 2019, 84, 334–348.
91. Jain, A.; Kanhangad, V. Investigating gender recognition in smartphones using accelerometer and gyroscope sensor readings. In Proceedings of the International Conference on Computational

Techniques in Information and Communication Technologies, New Delhi, India, 11–13 March 2016.

92. Meena, T.; Sarawadekar, K. Gender recognition using in-built inertial sensors of smartphone. In Proceedings of the IEEE Region 10 Conference, Hyderabad, India, 16–19 November 2020; pp. 462–467.
93. Singh, S.; Shila, D.M.; Kaiser, G. Side channel attack on smartphone sensors to infer gender of the user: Poster abstract. In Proceedings of the Conference on Embedded Networked Sensor Systems, New York, NY, USA, 10 November 2019; pp. 436–437.
94. Ngo, T.T.; Ahad, M.A.R.; Antar, A.D.; Ahmed, M.; Muramatsu, D.; Makihara, Y.; Yagi, Y.; Inoue, S.; Hossain, T.; Hattori, Y. OU-ISIR wearable sensor-based gait challenge: Age and gender. In Proceedings of the International Conference on Biometrics, Crete, Greece, 4–7 June 2019.
95. Sabir, A.; Maghdid, H.; Asaad, S.; Ahmed, M.; Asaad, A. Gait-based gender classification using smartphone accelerometer sensor. In Proceedings of the International Conference on Frontiers of Signal Processing, Marseille, France, 18–20 September 2019; pp. 12–20.
96. Acien, A.; Morales, A.; Fierrez, J.; Vera-Rodriguez, R.; Hernandez-Ortega, J. Active detection of age groups based on touch interaction. IET Biom. 2019, 8, 101–108.
97. Tolosana, R.; Ruiz-Garcia, J.C.; Vera-Rodriguez, R.; Herreros-Rodriguez, J.; Romero-Tapiador, S.; Morales, A.; Fierrez, J. Child-computer interaction: Recent works, new dataset, and age detection. arXiv 2021, arXiv:2102.01405.
98. Miguel-Hurtado, O.; Stevenage, S.; Bevan, C.; Guest, R. Predicting sex as a soft-biometrics from device interaction swipe gestures. Pattern Recognit. Lett. 2016, 79, 44–51.
99. Jain, A.; Kanhangad, V. Gender recognition in smartphones using touchscreen gestures. Pattern Recognit. Lett. 2019, 125, 604–611.
100. Almaatouq, A.; Prieto Castrillo, F.; Pentland, A. Mobile communication signatures of unemployment. In Proceedings of the International Conference on Social Informatics, Bellevue, WA, USA, 14–17 November 2016; pp. 407–418.
101. Yuan, Y.; Raubal, M.; Liu, Y. Correlating mobile phone usage and travel behavior—A case study of Harbin, China. Comput. Environ. Urban Syst. 2012, 36, 118–130.
102. Scherrer, L.; Tomko, M.; Ranacher, P.; Weibel, R. Travelers or locals? Identifying meaningful sub-populations from human movement data in the absence of ground truth. EPJ Data Sci. 2018, 7, 19.
103. Riederer, C.; Zimmeck, S.; Phanord, C.; Chaintreau, A.; Bellovin, S. I don’t have a photograph, but you can have my footprints. Revealing the demographics of location data. In Proceedings of

the ACM on Conference on Online Social Networks, Palo Alto, CA, USA, 2–3 November 2015; pp. 185–195.

104. Wu, L.; Yang, L.; Huang, Z.; Wang, Y.; Chai, Y.; Peng, X.; Liu, Y. Inferring demographics from human trajectories and geographical context. Comput. Environ. Urban Syst. 2019, 77, 101368.
105. The eXtreme Gradient Boosting Library. 2022. Available online: https://xgboost.ai/about (accessed on 4 November 2022).
106. Neal, T.; Woodard, D. A gender-specific behavioral analysis ofmobile device usage data. In Proceedings of the International Conference on Identity, Security, and Behavior Analysis, Singapore, 10–18 January 2018; pp. 1–8.
107. Chen, K.; Zhang, D.; Yao, L.; Guo, B.; Yu, Z.; Liu, Y. Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities. ACM Comput. Surv. 2021, 54, 77.
108. Sun, L.; Zhang, D.; Li, B.; Guo, B.; Li, S. Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations. Ubiquitous Intell. Comput. 2010, 6406, 548–562.
109. Thomaz, E.; Essa, I.; Abowd, G.D. A practical approach for recognizing eating moments with wrist-mounted inertial sensing. In Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing, Osaka, Japan, 7–11 September 2015; pp. 1029–1040.
110. Santani, D.; Do, T.; Labhart, F.; Landolt, S.; Kuntsche, E.; Gatica-Perez, D. DrinkSense: Characterizing youth drinking behavior using smartphones. IEEE Trans. Mob. Comput. 2018, 17, 2279–2292.
111. Arnold, Z.; Larose, D.; Agu, E. Smartphone inference of alcohol consumption levels from gait. In Proceedings of the 2015 International Conference on Healthcare Informatics, Dallas, TX, USA, 21–23 October 2015; pp. 417–426.
112. Chang, L.; Lu, J.; Wang, J.; Chen, X.; Fang, D.; Tang, Z.; Nurmi, P.; Wang, Z. SleepGuard: Capturing rich sleep information using smartwatch sensing data. In Proceedings of the 2015 ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies; ACM: New York, NY, USA, 2018; Volume 2, pp. 1–34.
113. Wan, N.; Lin, G. Classifying human activity patterns from smartphone collected GPS data: A fuzzy classification and aggregation approach. Trans. GIS 2016, 20, 869–886.
114. Chen, Z.; Zhang, L.; Jiang, C.; Cao, Z.; Cui, W. WiFi CSI based passive human activity recognition using attention based BLSTM. IEEE Trans. Mob. Comput. 2018, 18, 2714–2724.
115. Ma, Y.; Arshad, S.; Muniraju, S.; Torkildson, E.; Rantala, E.; Doppler, K.; Zhou, G. Location-and person-independent activity recognition with Wifi, deep neural networks, and reinforcement learning. ACM Trans. Internet Things 2021, 2, 1–25.

116. Yao, Y.; Song, L.; Ye, J. Motion-To-BMI: Using motion sensors to predict the body mass index of smartphone users. Sensors 2020, 20, 1134.
117. Albanese, E.; Launer, L.; Egger, M.; Prince, M.; Giannakopoulos, P.; Wolters, F.; Egan, K. Body mass index in midlife and dementia: Systematic review and meta-regression analysis of 589,649 men and women followed in longitudinal studies. Alzheimer's Dementia Diagn. Assess. Dis. Monit. 2017, 8, 165–178.
118. Dobner, J.; Kaser, S. Body mass index and the risk of infection-from underweight to obesity. Clin. Microbiol. Infect. 2018, 24, 24–28.
119. Garcia-Ceja, E.; Riegler, M.; Nordgreen, T.; Jakobsen, P.; Oedegaard, K.J.; Tørresen, J. Mental health monitoring with multimodal sensing andmachine learning: A survey. Pervasive Mob. Comput. 2018, 51, 1–26.
120. Arroyo-Gallego, T.; Ledesma-Carbayo, M.J.; Sanchez-Ferro, A.; Butterworth, I.; Mendoza, C.S.; Matarazzo, M.; Montero, P.; Lopez-Blanco, R.; Puertas-Martin, V.; Trincado, R.; et al. Detection of motor impairment in Parkinson's disease via mobile touchscreen typing. IEEE Trans. Biomed. Eng. 2017, 64, 1994–2002.
121. Castrillon, R.; Acien, A.; Orozco-Arroyave, J.R.; Morales, A.; Vargas, J.F.; Vera-Rodrıguez, R.; Fierrez, J.; Ortega-Garcia, J.; Villegas, A. Characterization of the handwriting skills as a biomarker for parkinson disease. In Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition (FG'19)–Human Health Monitoring Based on Computer Vision, Lille, France, 14–18 May 2019.
122. Bevan, C.; Fraser, D. Different strokes for different folks? Revealing the physical characteristics of smartphone users from their swipe gestures. Int. J. Hum. Comput. Stud. 2016, 88, 51–61.
123. Palmius, N.; Tsanas, A.; Saunders, K.; Bilderbeck, A.C.; Geddes, J.R.; Goodwin, G.M.; De Vos, M. Detecting bipolar depression from geographic location data. IEEE Trans. Biomed. Eng. 2016, 64, 1761–1771.
124. Tal, A.; Shinar, Z.; Shaki, D.; Codish, S.; Goldbart, A. Validation of contact-free sleep monitoring device with comparison to polysomnography. J. Clin. Sleep Med. 2017, 13, 517–522.
125. Behar, J.; Roebuck, A.; Shahid, M.; Daly, J.; Hallack, A.; Palmius, N.; Stradling, J.; Clifford, G.D. SleepAp: An automated obstructive sleep apnoea screening application for smartphones. IEEE J. Biomed. Health Inform. 2014, 19, 325–331.
126. Kostopoulos, P.; Nunes, T.; Salvi, K.; Togneri, M.; Deriaz, M. StayActive: An application for detecting stress. In Proceedings of the International Conference on Communications, Computation, Networks and Technologies, Barcelona, Spain, 15–20 November 2015.
127. Neal, T.; Canavan, S. Mood versus identity: Studying the iinfluence of affective states on mobile biometrics. In Proceedings of the IEEE International Conference on Automatic Face and Gesture,

Buenos Aires, Argentina, 16–20 November 2020.

128. Quiroz, J.C.; Geangu, E.; Yong, M.H. Emotion recognition using smart watch sensor data: Mixed-design study. JMIR Mental Health 2018, 5, e10153.
129. Cao, B.; Zheng, L.; Zhang, C.; Yu, P.; Piscitello, A.; Zulueta, J.; Ajilore, O.; Ryan, K.; Leow, A. DeepMood: Modeling mobile phone typing dynamics for mood detection. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017.
130. Hung, G.; Yang, P.; Chang, C.; Chiang, J.; Chen, Y. Predicting negative emotions based on mobile phone usage patterns: An exploratory study. JMIR Res. Protoc. 2016, 5, e160.
131. Gao, Y.; Bianchi-Berthouze, N.; Meng, H. What does touch tell us about emotions in touchscreen-based gameplay? ACM Trans. Comput. Hum. Interact. 2012, 19, 1–30.
132. Shah, S.; Teja, J.; Bhattacharya, S. Towards affective touch interaction: Predicting mobile user emotion from finger strokes. J. Interact. Sci. 2015, 3, 6.
133. Zhang, X.; Li, W.; Chen, X.; Lu, S. MoodExplorer: Towards compound emotion detection via smartphone sensing. Proc. Acm Interactive Mobile Wearable Ubiquitous Technol. 2018, 1, 1–30.
134. Nguyen, K.A.; Akram, R.N.; Markantonakis, K.; Luo, Z.; Watkins, C. Location tracking using smartphone accelerometer and magnetometer traces. In Proceedings of the International Conference on Availability, Reliability and Security, University of Kent, Canterbury, UK, 26–29 August 2019.
135. Hua, J.; Shen, Z.; Zhong, S. We can track you if you take the metro: Tracking metro riders using accelerometers on smartphones. IEEE Trans. Inf. Forensics Secur. 2017, 12, 286–297.
136. Han, J.; Owusu, E.; Nguyen, L.T.; Perrig, A.; Zhang, J. ACComplice: Location inference using accelerometers on smartphones. In Proceedings of the 4th International Conference on Communication Systems and Networks, Rajkot, Gujrat, India, 11–13 May 2012.
137. Singh, V.; Aggarwal, G.; Ujwal, B.V.S. Ensemble based real-time indoor localization using stray Wifi signal. In Proceedings of the IEEE International Conference on Consumer Electronics (ICCE'18), Las Vegas, NV, USA, 12–15 January 2018; pp. 1–5.
138. Cai, L.; Chen, H. TouchLogger: Inferring keystrokes on touch screen from smartphone motion. HotSec 2011, 11, 9.
139. Owusu, E.; Han, J.; Das, S.; Perrig, A.; Zhang, J. ACCessory: Password inference using accelerometers on smartphones. In Proceedings of the Workshop on Mobile Computing Systems and Applications, San Diego, CA, USA, 28–29 February 2012.
140. Aviv, A.J.; Sapp, B.; Blaze, M.; Smith, J.M. Practicality of accelerometer side channels on smartphones. In Proceedings of the Annual Computer Security Applications Conference, Orlando,

FL, USA, 3–7 December 2012.