SoK: Your Mind Tells a Lot About You: On the Privacy Leakage via Brainwave Devices

Anuradha Mandal
University of Alabama at Birmingham
Birmingham, Alabama, USA
anuradha@uab.edu

Nitesh Saxena
Texas A&M University
College Station, Texas, USA
nsaxena@tamu.edu

ABSTRACT

Head-worn wearables, such as consumer-grade EEG headsets deployed in Brain Computer Interfaces (BCI), are getting popularity in the gaming and entertainment industry, and for people with certain disabilities. However, the increasing popularity of these wearables creates a significant privacy risk. For instance, tech companies are intending to use brainwave devices to detect workers' emotional state and mental condition. There are AI techniques that can learn what people are looking at in real-time. Silently conversing with the computing system is now possible using neuromuscular signals, for instance, untold digit recognition with higher accuracy is possible, which can retrieve untold PIN or password. These applications can reveal more private information than designated benign purpose, such as, while detecting performance of worker, sensitive information like Parkinson's disease, substance abuse disorder, heart disease, can be revealed from brainwave. The consequences of these privacy leakages may be potentially devastating, such as tracking users for targeted advertisements and launching targeted attacks against users.

In this paper, we analyze current devices, explore previously studied attacks, research efforts to extract information from brainwave and analyze and synthesize potential future attacks from the current deployment. This systematization will provide right direction towards ensuring privacy risk of BCI devices, which is a pre-requisite to building future defense mechanisms against the attacks.

CCS CONCEPTS

• **Security and privacy**;

KEYWORDS

Side-channel attacks, Brainwave security & privacy, Wireless or mobile security for cyber-physical systems, SoK

ACM Reference Format:
Anuradha Mandal and Nitesh Saxena. 2022. SoK: Your Mind Tells a Lot About You: On the Privacy Leakage via Brainwave Devices. In *Proceedings of the 15th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '22), May 16–19, 2022, San Antonio, TX, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3507657.3528541

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
WiSec '22, May 16–19, 2022, San Antonio, TX, USA
© 2022 Association for Computing Machinery.
ACM ISBN 978-1-4503-9216-7/22/05…$15.00
https://doi.org/10.1145/3507657.3528541

1 INTRODUCTION

Wearable devices are getting extensively deployed in activity tracking and monitoring in day-to-day life. These devices are examples of Internet-of-Things which come in the form of electronic sensors that enable objects to exchange data through the Internet with another electronic device without human intervention. The BCI headsets are getting popularity in the gaming and entertainment industries. Consumer-grade BCI devices are easily accessible and are being used in a variety of applications, such as relaxation training, video games and hand-free keyboard.

Scientists are working extensively on BCI technology to offer a direct link between the gray matter of the human brain and the computing system. The researchers are working to make human operate-computer using their brains [8, 13]. Mike Ambinder, a psychologist in Valve, mentioned [15] gaming industry approaches of naturalistic method to control the gaming environment which could improve the connection between the virtual and real-world instead of using a typical 17-button controller [3]. While the gameplay may need eye-tracking data, sweat level, physiological state to convey command to the computing system, but the BCI headset being used in gameplay can measure extra things like heart rate, facial expression, body temperature etc.

With the increasing popularity of BCI devices, the privacy leakage linked with the sensors have become a research area to be explored. For example, some tech companies are mining data from workers brain via BCI devices in an industrial setting to monitor the emotional state of workers to increase productivity, in such scenario privacy can be compromised very easily by listening onto brainwaves and extracting private information from it (i.e., emotional state) [22]. A trial of BCI devices used in a school in China to monitor pupils' brainwaves has been halted due to the privacy concern of the pupils [81]. An AI has been created that can draw what a person is looking at in real-time by reading and extracting brainwave via EEG headset [32]. Malicious BCI application can take advantage of benign purpose by recording brainwaves, analyzing data and extracting features to classify and get private information from app-users' brain without the consent or awareness [41]. For relaxation training, BCI headsets are widely used and this can potentially leak private information using brainwaves.

Moreover, brainwave signals can be accessed by smartphone applications without users' permission [21], which can continuously record brainwaves and leak sensitive information. Malicious applications can continuously record the users' activities passively using brainwaves, and they can release this information to malicious

third parties for future attacks. In the future, BCI devices might lead technology to offer fast and technologically advanced life. This type of deployment of BCI and availability to end users and the future BCI deployment, raise privacy concern for the end users. Figure 1 represents the higher-level overview of threat: EEG signal is being recorded by a malicious app or passively being recorded by the attacker from smartphone, computer, javascript enabled website. Afterward, the attacker use the signal to extract features using its own classification and learn private sensitive information about the victim. Private information could be mental and emotional state, medical condition, alcoholism, age group, PIN, password etc. which can later be useful for malicious purpose, i.e., targeted advertisements, selling data to third party, etc.

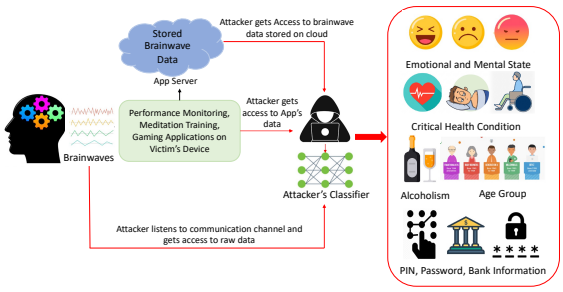


Figure 1: Overview of Brainwave Attacks

In this systematization, we surveyed research papers and systematized the privacy leakage via consumer-grade EEG headsets by studying the existing method to read neural signals, existing passive and active attacks and benign BCI enabled applications and research works which can reveal private information about user when malicious party gets access to data. Briefly, the contributions of this paper are as follows:

(1) We provide a brief overview of consumer-grade brain mapping technique, components of BCI system in order to understand the threat related to BCI (Sec. 2) and available BCI devices (Sec. 3).
(2) We present a threat model of a compromised BCI system which can steal brainwaves (Sec. 4).
(3) We provide a detailed review and systemization of studied passive (Sec. 5.1) and active (Sec. 5.2) attacks to extract private information from brainwaves.
(4) Next, we present a systemization of existing works which demonstrates extraction of information from brainwaves for benign purpose, later which can be used as ground truth for extracting private sensitive information for malicious purpose (Sec. 5.3).
(5) We present a brief systemization of BCI use in different sectors and the privacy implication related to it (Sec. 6).
(6) Lastly, we present and evaluate existing research efforts on privacy enhancement to protect sensitive brainwave-information (Sec.7).

2 BACKGROUND

As a background of this paper, we discuss about brain mapping technique being used in commercial industry, brain regions' functionalities and information gathered from them, components of a BCI system.

2.1 Consumer-Grade Brain Computer Interface

Brain mapping techniques are getting popularity in gaming and entertainment industry, as a form of VR, as a portable headset to give commands to games, to monitor students' attention level, to detect meditation status during yoga. There are several consumer-grade brain computer interfaces available on market today. These BCI devices mainly use Electroencephalography (EEG) signal to record brainwave. EEG devices are typically noninvasive, with electrode placed along with scalp. Clinically EEG is used to determine changes in brain activity especially epilepsy or another seizure disorder which helps to identify brain tumor, brain damages from head injury, inflammation of the brain (encephalitis), stroke, sleep disorders etc. EEG is used extensively in neuroscience, cognitive science, cognitive psychology, neurolinguistics, psychophysiological and computer science research.

2.2 Functionalities of Different Regions of Human Brain

In this section, we discuss about brief of human brain's different regions' functionalities which will give a good direction to understand what positioning of sensors can gather what type of information. Discussion below gives a high-level overview of what type of data can be learnt from different regions of brain using neuro-imaging techniques. Table 1 is a brief snapshot of this discussion.

Function of Forebrain: The Cerebrum, also known as the cerebral cortex, the largest part of the human brain is associated with higher brain function such as thought and action. The Cerebrum is divided into two cerebral hemispheres and each hemisphere is conventionally divided into four lobes – the frontal, temporal, parietal, and occipital lobes [5]. The **frontal lobe** is associated with brain's ability to reason, organize, plan, speak, move, make facial expressions, serial task, problem solving, control inhibition, spontaneity, initiate and self-regulate behaviors, pay attention, remember and control emotions [6]. The **parietal lobe** controls our complex behaviors, including senses, such as vision, touch, body awareness and spatial orientation, integrates sensory information from various parts of our body, visuospatial processing, language comprehension, ability to construct, body positioning and movement, neglect/inattention, left-right differentiation and self-awareness/insight [12]. The **occipital lobe** is associated with our visual processing, such as visual recognition, visual attention, spatial analysis (moving in a 3-D world) and visual perception of body language; such as postures, expressions and gestures [11]. The **temporal lobe** is associated with processing our perception and recognition of auditory stimuli (including our ability to focus on one sound among many, like listening to one voice among many at a party), comprehending spoken language, verbal memory, visual memory and language production (including fluency and word-finding), general knowledge and autobiographical memories [14].

Function of Midbrain: The midbrain is located below the cerebral cortex, and the primary role of the midbrain is to act as a sort of relay station for our visual and auditory systems [7].

Function of Hindbrain: The cerebellum is a major structure of the hindbrain that is located near brainstem which is responsible for coordinating voluntary movements and functions including

Table 1: Brain Regions' Functionalities and Information Processing

Regions of Brain	lobes	Functionality
Cerebral Cortex	Frontal Lobe [6]	Reasoning, organize, plan, speak, move, make facial expressions, serial task, problem solving, control inhibition, spontaneity, initiate and self-regulate behaviors, pay attention, remember and control emotions
	Parietal Lobe [12]	Complex behaviors like vision, touch, body awareness and spatial orientation, body positioning and movement, neglect/inattention, left-right differentiation and self-awareness/insight etc.
	Occipital Lobe [11]	Visual processing like visual recognition, visual attention, spatial analysis (moving in a 3-D world) and visual perception of body language; such as postures, expressions and gestures
	Temporal Lobe [14]	Processing perception and recognition of auditory stimuli
Subcortical Region	Cerebrum and Cerebellum [1]	Digestion, breathing, heart rate, and information transfer from the cerebrum to cerebellum

motor skills such as balance, coordination, posture (eye movements and movements associated with speaking), mental function (thinking, language processing and mood, attention, fear response, and pleasure or reward response), body balance and posture (walking, standing), motor learning (riding a bike or hitting a baseball), vision [4].

2.3 Components of BCI System

A BCI system consists of the following components: signal acquisition, preprocessing, feature extraction, classification and application interfaces [69]. Figure 2 demonstrates components of a standard BCI system.

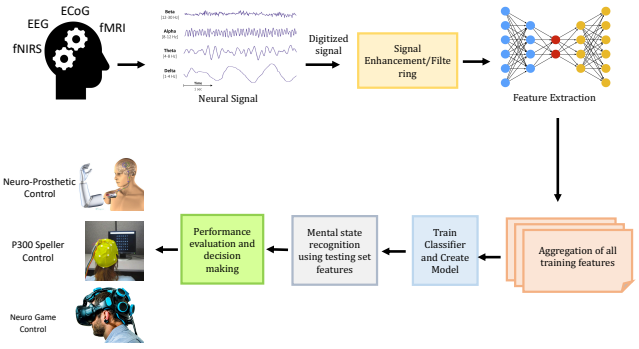


Figure 2: Components of a BCI systems.

Signal Acquisition: Signal acquisition is the measurement of brain signals using a particular sensor modality (i.e., scalp or intracranial electrodes for electrophysiologic activity). The signals are amplified to levels suitable for electronic processing. The signals are then digitized and transmitted to a computing system.

Feature Extraction: Feature extraction is the process of analyzing the digital signals to distinguish pertinent signal characteristics (i.e., signal features related to the person's intent) from extraneous content and representing them in a compact form suitable for translation into output commands. The most commonly extracted signal features in current BCI systems are time-triggered EEG response amplitudes and latencies, power within specific EEG frequency bands, or firing rates of individual cortical neurons. Environmental artifacts and physiologic artifacts such as electromyographic signals are avoided or removed to ensure accurate measurement of the brain signal features.

Feature Translation: The resulting signal features are then passed to the feature translation algorithm, which converts the features into the appropriate commands for the output device (i.e., commands that accomplish the user's intent). The translation algorithm should be dynamic to accommodate and adapt to spontaneous or learned changes in the signal features and to ensure that the user's possible range of feature values covers the full range of device control.

Application Interfaces: The commands from the feature translation algorithm operate the external device, providing functions such as letter selection, cursor control, robotic arm operation, and decision making.

3 INSTANCES OF CURRENT BCI DEVICES

In this section, we discuss the current BCI devices available in the market and are being analyzed for better results. Table 2 represents a list of BCI devices and their functionalities.

3.1 EEG devices in Commercial Use:

Neurosky Mindwave Mobile EEG headsets are result of EEG biosensor technology research which is portable and easy-to-control and all in one wearable package. This device works by monitoring these electrical impulses with a forehead sensor FP1, record signal at 512Hz. It measures electrical signals and calculated interpretations are then output as digital messages to the computing system. This headset is being used for educational training, attention and meditation measurements, gaming and entertainment etc. [9].

B-Alert X-Series device is being used for operational neuroscience applications and cognitive state assessments which allows high quality data to be acquired in real or virtual environments by personnel with limited technical training [2]. The X-10 version has 9 high-quality EEG channels (F3, F4, FOz, P3, P4, POz, C3, C4, COz) and 1 optional channel for ECG, EMG, or EOG which records signal at 256Hz. B-alert X-24 has 20 EEG channels in sensor positioning to Fz, F1, F2, F3, F4, Cz, C1, C2, C3, C4, CPz, Pz, P1, P2, P3, P4, POz, Oz, O1, O2.

Emotive Headsets monitor brain activity. Its performance metrics provide real-time detection of cognitive states, so users can get valuable insights from the headset right away [28]. The Emotive Epoc uses 14 EEG channels (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4) and Emotive Insight uses 5 EEG channels (AF3, AF4, T7, T8, Pz) with a sampling rate of 2048 Hz (internal) to collect brain signal for performance measurements.

NextMind is a visual EEG headset with a dry electrode which can be combined with headbands, hats, and AR/VR headsets. This headset allows users to control the visual interface in real-world [10].

Table 2: Instances of Current BCI, Sensor Placements and Tasks

Headsets	Sensors	Tasks
NeuroSky [9]	EEG: FP1	Educational training, gaming and entertainments.
Emotive Headsets [28]	Emotive Epoc (EEG: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4), Emotive Insight (EEG: AF3, AF4, T7, T8, Pz)	Records cognitive states
NextMind [10]	9 electrodes	Controls visual interfaces
B-Alert X-Series [2]	B-Alert X10 (EEG: F3, F4, FOz, P3, P4, POz, C3, C4, COz), B-Alert X24 (EEG: Fz, F1, F2, F3, F4, Cz, C1, C2, C3, C4, CPz, Pz, P1, P2, P3, P4, POz, Oz, O1, O2)	Records cognitive state/human factors assessments
Thinking Cap [38]	EEG sensors and Bluetooth speaker	Measure the self-esteem of children
AlterEgo [36]	Bone conducted headphone	Detects silent speech
AttentivU [39]	EEG, EOG	Cognitive load, fatigue, engagement, and focus (EEG), eye movements (EOG)

3.2 EEG devices on Research Lab:

AlterEgo, a wearable interface that allows user to silently converse with a computing device without any voice or any discernible movements can compromise users' privacy by neuromuscular signals in internal speech articulators [36]. Unlike head-worn wearables, this interface uses face and neck area for silent speech signals and accuracy of digit recognition is 92% which may lead to privacy leakage when someone recalls PINs and passwords.

Thinking Cap is a wearable system that communicates praise for effort and ability in order to improve the resilience and self-esteem of the student wearing it to positively influence their motivation and academic achievements (momentary learning). It is built into a "Sorting Hat" from the Harry Potter franchise, with an embedded electroencephalography (EEG) headset and a Bluetooth speaker which can recognize several mental processes like motor, auditory, or visual imagery as well as cognitive load and engagement level of the child [38].

AttentivU uses both EEG and EOG for real-time monitoring of physiological data. The device is designed as a socially acceptable pair of glasses and employs silver electrodes [39].

4 METHODS OF EXTRACTING PRIVATE INFO FROM A COMPROMISED BCI SYSTEM

There are different ways to run attack on a brainwave signal. For example, an attacker can get access to a signal from cloud storage where the signal is being stored or directly by hijacking communication channel. An attacker can also launch an attack remotely using javascript while the victim is connected to a malicious website. Moreover, EEG signals can be collected using Bluetooth from a neural-based application (e.g., a smartphone application) without user intervention. A malicious app could take this advantage to get victim's neural signal and later use it for malicious purposes.

Bonaci et al. [21] mentioned existing BCI open-development platforms typically grant every application developer full control over all components. The threat model discussed in the paper, assumed an attacker has access to all of the resources: acquisition system, application, signal processing system, feature extraction, decoding algorithm. The first type of attacker extracts users' private information by hijacking the legitimate components of a BCI system. Such attacker exploits for malicious purposes those feature extraction and decoding algorithms that are intended for the legitimate BCI applications. The second type of attacker extracts users' private information by adding or replacing the legitimate BCI components. Such attacker implements additional feature extraction and decoding algorithms, and either replaces or supplements the existing BCI components with the additional malicious code.

From Figure 3, we can observe the difference between two types of attacker is only in the structure of the "brain malware" component. To steal private information through brainwaves, the attacker interacts with users by presenting them with specific sets of stimuli and recording their responses to the presented stimuli. There are several well-established methods of presenting stimuli to users: Oddball paradigm - a technique where users are asked to react to specific stimuli, referred to as target stimuli, hidden as rare occurrences in a sequence of more common, non-target stimuli [36]. Guilty knowledge test - a technique based on the hypothesis that a familiar stimulus evokes a different response when viewed in the context of similar, but unfamiliar items [66]. Priming - a technique that uses an implicit memory effect where one stimulus may have an influence on a person's response to a later stimulus [73].

5 SYSTEMATIZATION AND EVALUATION OF BCI ATTACKS

From this section, we learn about existing passive and active attacks which utilized different attack scenarios to learn about user's PIN, password, age group, alcoholism, face recognition etc.

5.1 Previously Studied Passive Attacks

The attack by Neupane et al., PEEP [81] was a passive attack which can passively eavesdrop and can get access to victim's brainwave while victim is typing sensitive private information like PINs, password on keyboard/keypad. PEEP is highly surreptitious because PEEP requires passive monitoring of brain signals, not deliberate, and active strategies which can trigger suspicion and can be detected by the user. Also, PEEP achieves orders of magnitude higher accuracies compared to prior active PIN inferring attacks. When a user may enter passwords or private credentials to their computers or mobile phones, while the BCI device is being worn by the user. At this point, a malicious application captures EEG signals when users are typing passwords or PINs in virtual or physical keyboards. Virtual keyboard PIN, virtual ATM PIN, physical numeric keypad PIN entry, physical keyboard password entry, these four different scenarios were tested with different parameters such as EEG devices (Emotiv vs. B-Alert), keypads (virtual vs. real), and data types (4-digit pin vs. 6-character password). Virtual Keyboard PIN Entry (VKPE) Attack measures the visual and metal of processing digits while a user is typing. In the VKPE attack user enters a 4-digit PIN codes in the text box using the mouse, fewer numbers than the previous scenario. In this attack scenario, the participants were asked to enter a 4-digit PIN code in a text box using a mouse.

The virtual key was flashed for 500ms or till the next key was pressed on the keyboard. This procedure was followed to make sure the participants are clicking on the right digit. In this attack, the

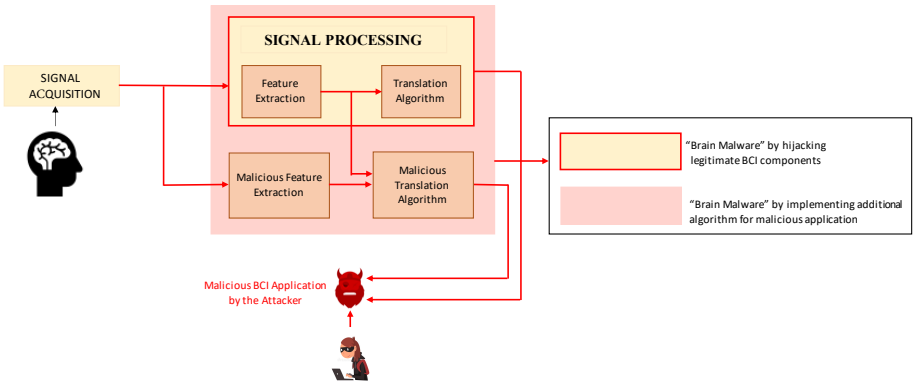


Figure 3: A simplified diagram of a compromised BCI system. Type 1 Attacker: exploits the legitimate feature extraction and decoding algorithms, and Type 2 Attacker: implements additions algorithms for malicious applications and either replaces or supplements the legitimate BCI resources

Table 3: List of Attack literatures, Classifiers Used to Extract Data and Retrieved Private Information from Brainwave

Studied Attacks	Attacker's Intuition	Classifiers	Success Rate
Feasibility of Side-Channel Attacks [56]	Learn 4-digit PIN, bank information, birth month, residence area	Boosted logistic regression, Stepwise Linear Discriminant Analysis, Fisher's linear discriminant analysis (LDA), Stepwise Linear Discriminant Analysis (SWLDA)	10%(PIN), 30%(bank information), 43%(Birth month), 30%(residence)
Subliminal brain activity [29]	Identifies known face and unknown face from an extremely short period of time	Band pass filtering, Machine learning classification	20.84%
PEEP [81]	Detects 4-digit PIN (VKPE, VAPE, PNKPE), 6-character password (PKPE)	Instance Based Learning (IBL), KStar algorithm, Naive Bayes (NB)	43% (VKPE), 33% (VAPE), 46% (PNKPE), 35% (PKPE)
Brain Hemorrhage [62]	Learn about Age Group and Alcoholic Behavior	Random Tree, Logistic Regression, Multilayer Perceptron, Support Vector, knearest neighbor algorithm Machines(SMO)	94% (age group) and 96% (alcoholic behavior)

attacker measured visual and mental processing of digits, eye and hand movement while the victim was typing the PIN. In Virtual ATM PIN Entry (VAPE), the attacker used a virtual ATM keyboard to reduce the number of keys on the keyboard. The attacker measured the similar parameters as in VKPE in this attack. The lower number of keys resulted in higher accuracy than VKPE. Physical Numeric Keypad PIN Entry (PNKPE) attack measured the mental processing along with the facial muscle, eye, hand, finger while typing 4-digit PIN on a physical keyboard unlike virtual keyboard. Unlike the previous two scenarios, the attacker used a numeric physical keyboard to create digit specific pattern in event-related potentials. Physical Keyboard Password Entry (PKPE) attack used a 6-character based passwords and measures facial muscle, eye, hand, finger while the user is typing. In this experiment, the participants were asked to enter a 6-character based passwords using a physical keyboard, like a laptop keyboard. The success rate of randomly guessing a digit of the PIN is 100/10 (10%) and the success rate of randomly guessing a character is 100/26 (3.84%). PEEP increases this accuracy of correctly identifying the digits of PIN to 47.5% and passwords to 34.7%. In Table 3, we summarize the attacks and the private data extracted from these attacks and in Appendix A, we can see the visual representation of the attack scenario using keyboard/keypad.

5.2 Previously Studied Active Attacks

Martinovic et al. [56] used ERPs as a vector of side-channel attack to snoop into users private information. The authors showed images of numbers, banks, and ATMs to the participants when their brain signals were measured. They used the brain signal to decrease entropy of information related to PIN, banks, ATMs by 23-40%. In this attack, the user was asked to memorize a 4-digit PIN at the beginning of the experiment to calibrate successfully. To extract bank information, the attacker showed a list of bank logos during the training phase and during the experiment the attacker showed list of credit cards related to the banks that were shown at the beginning. Based on the neural pattern during the experiment and the calibration phase, the attacker predicted the bank information of the user. To extract the known location or home address of the victim, the attacker designed a malicious app where the user was presented with a list of location highlighted map. Based on which location on the map triggered the neural pattern, the attacker could guess the location of the user's known or living place. Another attack was designed by Frank et al in [29] to detect subliminal brain activity. This subliminal attack was performed for less than 13.3 milliseconds in which the visual probing was tested. The classifier of this attack can extract information from the recorded EEG signal to identify brain activity related to a known face that the user subliminally recognizes, an unknown face and a plain video sequence without any subliminal stimulation. A study conducted in [62] revealed user's age group and alcoholic behavior and success rate of the classifier is approximately 94% and 96% respectively, due to higher end BCI datasets. The Hemorrhage attack was designed using machine learning techniques to identify victims' age group and alcoholic behavior while the victim is watching videos or viewing images. Due to the higher-end medical grade BCI device, this attack reached higher accuracy from machine learning classifiers. Appendix B represents the stimuli used in the attacks and the success rate of different attacks.

Table 4: List of works that Extracted Private Data from Brainwave for Benign Purpose, Feature Extraction Methods, Machine Learning Classifiers Used to Extract Information, Extracted Information from Brainwave Signal and Accuracy of Classifiers

Literatures	Feature Extraction	ML Classification	Extracted Information	Accuracy
Liu et al. [20, 52]	Band-Pass Filtering	Fractal Dimension (FD)	Emotion: sad, frustrated, fear, satisfied, pleasant and happy	84.9%
Liu et al. [49]	ResNets, LFCC	KNN, SVM, LR, RF, NB, DT and FC	Emotion: anger, joy, sadness and pleasure	KNN: 89.72%
Zheng et al. [82]	DE, DASM, RASM	DBN, SVM, LR and KNN	Emotion: positive, neutral and negative	DBN: 86.08%, SVM: 83.99%, LR: 82.70%, KNN: 72.60%
Correia et al. [25]	Temporal-windows, Time-frequency, MVPA	linear-SVM	Spoken word detection in bilingual	98.3%
Soman et al. [75]	Short-Time Fourier Transform (STFT)	SVM, LDA, Gaussian	Language discrimination: English vs Japanese and Hindi vs Japanese	E-J-SVM:64.06, LDA:62.79, Gaussian:58.64; H-J-SVM: 62.57%, LDA: 52.18%, Gaussian: 65.09%
Krishna et al. [40]	Connectionist Temporal Classification (CTC), Attention based RNN, RNN transducer model	GAN, WGAN, LSTM Regression	Word and Character Recognition	RNN-T(WER): 92.98 %, 69.89 %, 70.37 %, 92.66 %, CTC(WER): 73.6%, 83.8 %, 91.1%, 91.5 %
Liu et al. [50]	FFT, PSD	SVM	Recognize Attention Level	90.64%
Wang et al. [80]	FDA, Statistical Features	SVM with gradient descent	Decoding English Alphabet Letters	46.61%
Herff et al. [30]	Elliptic IIR low-pass and high-pass filters		Text Identification	>50%
Dan et al. [63]	FFT	SVM	Emotion: negative and positive	87.53%
Zheng et al. [26, 31, 78, 83]	PSD, DE, DASM, RASM, ASM, DCAU	KNN, LR, SVM, GELM	Emotion: negative, positive and neutral	KNN: 70.43%, LR: 84.08%, SVM: 78.21%, GELM: 91.07%
Li et al. [44]	CSP	linear-SVM	Happiness and Sadness	93.5%
Murugappan et al. [60]	Wavelet transform	KNN and LDA	Emotion: disgust, happy, surprise, fear and neutral	KNN: 77.68% LDA: 73.5%
Wang et al. [79]	Wavelet transform, PCA, LDA, CFS	linear-SVM	Negative and Positive Impression	87.53%
Petrantonakis et al. [61, 64, 76]	Statistical values, wavelet transform and HOC	QDA, KNN, MD and SVMs	Happiness, surprise, anger, fear, disgust and sadness	QDA: 62.3%, SVM: 83.33%, MD: 44.90%, KNN: 34.60%
Duan et al. [27]	DE, DASM, RASM and ES	linear-SVM and kNN	Negative and Positive Emotion	SVM: 74.10%, KNN: 69.24%
Pfurtscheller et al. [65]	DSLVQ	KNN	Left or Right Hand Movement	80%
Liu et al. [51]	Statistical Features	SVM	Left or Right Hand Movement	89.17%
Vanitha et al. [77]	Hilbert Huang Transform (HHT)	SVM, LDA, QDA, KNN	Stress Level Detection	SVM: 89.07%, LDA: 70.17%, QDA: 76.83%, KNN: 72.67%
Saeed et al. [71]	Power Spectral Densities	KNN, NB, SVM, LR, MLP	Long-Term Stress Detection	SVM, LR: 85.20%
Purnamasari et al. [67]	FFT	KNN	Stress and Meditation Level	80%
Jebelli et al. [33]	TDA, FDA	SVM	Stress Recognition	80.32%
Ji et al. [34]	EEG Bands	Deep Learning	Stress Index	90.96%
Amin et al. [18]	Wavelet Analysis, FDR, PCA	KNN, SVM, MLP, NB	Pattern Recognition	KNN: 93.33%
Liang et al. [47]	Burg AR model, Linear discriminant analysis	KNN, SVM, NMF, PCA, ANN	Identity Recognition	98.12%
Bird et al. [19]	FFT, Statistical Features	Naive Bayes, Bayes Net, J48, Random Tree, Random Forest, MLP, SVM	Mental State Recognition	Bayes Net: 73.67%, J48: 80.65%, Random Tree: 76.21%, Random Forest: 87.16%, MLP: 80.85%, SVM: 75.24%
Makin et al. [54]	Temporal Convolutional Filters	RNN with LSTM	Cortical Activity to Text	Word Error Rate <3%

5.3 Studied Works to Extract Private Information from Neural Pattern

There are many studies conducted to extract private information from human brain, such as emotional state of mind, classify stress level, hand movement, identity recognition, pattern recognition etc. In this section, we discuss about major studies conducted to extract data from neural pattern and analyzed data with different classification methods to reach higher accuracy rates. Table 4 shows the briefs of the studied works, features extraction methods, machine learning classifiers to classify pattern of task and accuracy of classifier in identifying patterns from neural signals. From this section we learn what type of information can be learnt in certain accuracy, which can lead to privacy threat if the brainwave is accessible by malicious user.

Liu et al. explored EEG-based motor imagery in [51] which is very useful in brain-computer interface. Electroencephalography (EEG) microstates reflect the spatial configuration of quasi-stable electrical potential topographies. Different microstates represent different brain functions. In this paper, microstate method was used to process the EEG-based motor imagery to obtain microstate. The single-trial EEG microstate sequences differences between two motor imagery tasks – imagination of left and right hand movement were investigated.

Liao et al. conducted a study to decode individual finger movements from one hand from EEG signal [48]. The support vector machine (SVM) classifier's accuracy in detecting finger movements based on movement-related spectral changes as features is 77.11% in average over all subjects in the experiment. Using the same classifier, they found 91.28% accuracy in three epilepsy patients using ECoG data. The authors claimed, the accuracy obtained from the classifier for EEG, ECoG are significantly higher than empirical guessing level (51.26%).

A survey by Li et al. listed different research efforts to classifies human emotion (Fear, Anger, Sadness, Joy, Surprise, Disgust) using various classifiers. Li et al. used linear support vector machines (LINEAR-SVM) to extract emotion from data of 62-channels with accuracy of 93.5% [44]. A study by Murugappan et al. using 64-channels EEG data, showed that KNN classifier has 79% accuracy in extracting emotional state [60]. Wang et al. used 62 channel brainwave data to classify emotional state. SVM with linear nuclei showed accuracy of 87.53% [79]. Method by Petrantonakis et al [64] showed that best classifier (SVM) obtained 83.33% success rate in average. Duan et al found 74% accuracy from SVM classifier and found that MRMR algorithm can effectively improve the accuracy

of classifier [27]. Other research efforts to classify emotional state is listed in Table 4.

EEG allows studying non-invasively with high temporal resolution neural dynamics of speech processing. The temporal dynamics of EEG signals are informative of temporal order effects during speech processing. The study of Correia et al [75] demonstrated the feasibility of MVPA to decode individual spoken words from EEG responses and to assess the spectro-temporal dynamics of their language invariant semantic-conceptual representations.

Zhang et al. [46] conducted a study to show the progress of brain science to identify emotion classification. In this study, different machine learning classifications were used to identify and classify state of consciousness in human brain in terms of frequency bands and psychological states.

There are studies that have suggested that it is feasible to recognize isolated aspects of speech from neural signals, such as auditory features, phones or one of a few isolated words. Herff et al. [30] conducted a study where continuously spoken speech can be decoded into the expressed words from intracranial electrocorticographic recordings. Specifically, they implemented a system, called Brain-To-Text which models single phones, employs techniques from automatic speech recognition (ASR), and thereby transforms brain activity while speaking into the corresponding textual representation. The result demonstrate that the system can achieve word error rates as low as 25% and phone error rates below 50%. Additionally, this approach contributes to the current understanding of the neural basis of continuous speech production by identifying those cortical regions that hold substantial information about individual phones.

Work by Vanitha et al. [77] demonstrated real time stress detection is possible using the EEG signal. To determine the stress level, they used several stimuli from different categories, such as Interpersonal stimuli: quarrel with friends and parents, split up with partner, crisis in family, conflict with room mate; Intrapersonal stimuli: public speech, financial constraints, personal health issues; Academic Workload: meeting deadlines, poor performance, inadequate resources, fear of failure, poor time management, Unclear contents, competition with peers. Based on the different categorical stimuli, the work showed that accuracy rate of SVM is 89.07%, accuracy rate of LDA is 70.17%, QDA is 76.83% and KNN is 72.67% and concluded that SVM found higher accuracy than other ML classifiers.

Makin et al. [54] showed how to decode brain signal to extract speech with higher accuracy using the electrocorticogram. In this study, they trained a recurrent neural network to encode each sentence-length sequence of neural activity into an abstract representation, and afterwards, decoded this representation, word by word and turn into an English sentence. For this experiment, each participant's data consists of spoken repeated 30-50 sentences, while the brain activity was recorded using 250 electrodes distributed over peri-Sylvian cortices. For data analysis, high-Gamma signal was extracted at 200Hz and clipped to spoken sentence length to supply the signal to an artificial neural network. Figure 4 represents the higher level overview of the study. The green window on the high-gamma EEG signal in purple is a single sample of a feature sequence. Each filter maps data from 12 sample wide window from all electrode to a single samples of feature, afterwards slides by 12 sample to generate next feature sequence, and process goes on until 100 feature sequences are generated. Afterwards, the input feature sequences are passed to the RNN encoder as demonstrated in Figure 4, which learns to summarize in a single hidden state. Hidden state of final encoder then initiated decoder RNN, which learns to predict the next word. This encoder-decoder has lower error (less than 3%).

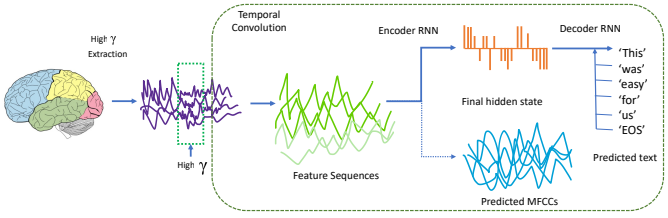


Figure 4: Overview of the Speech Encoder-Decoder from Electrocorticogram [54].

There are more works listed in Table 4 which could extract pattern to recognize sensitive/private information from human brainwave. From the table, we find, different feature extraction algorithms and machine learning classifiers were used to recognize pattern from brainwave.

6 OTHER LIKELY ATTACKS

In this section, we synthesize existing studies conducted to extract information from the human brain for benign purpose in different sectors and how accurately machine learning classifiers performed to extract information. Table 5 summarizes the discussion about likely privacy threats in different sectors. We also discuss about the possible privacy threats can be launched by attacker in similar situation. Table 6 represents practicality of attack scenarios in different sectors on BCI implementation and potentiality of the attacks.

6.1 Privacy Threat of Neuro-Medical Information

Since BCI has enabled people to communicate with the computer using neural sensor without users' intervention, it is widely used in medical area to help patients, especially with neurological disorders. A system proposed by Sharanreddy et al. [74] could recognize EEG abnormalities to detect brain tumors and epilepsy seizures. The applications developed within this field range from the control of prosthetic limbs and wheelchairs to the use in brain stimulation procedures [43]. Neurofeedback has been used for various treatments [57] like attention-deficit hyperactivity disorder (ADHD) [53]. Neurofeedback involves EEG activity recording and providing feedback with presence of a predetermined EEG features [58]. MacFarland et al. mentioned, Parkinson's disease and motor imaginary can be identified from BCI devices [57].

If the medical system get compromised, attackers can use brainwave data for malicious purpose. In such environment, the attacker can hack the server or can eavesdrop to transmitted channel to collect data and then decompose the raw signal to get private information about patients. To launch such attack, the attacker does not need to be physically present in the proximity of the victim.

Table 5: Privacy Implications of Benign BCI Applications Being Used in Different Sectors (Neuro-Medical, Silent Speech Detection, Speech Recognition, Neural Authentication, Gaming & Entertainment)

BCI Enabled Applications	Applications' Jobs	Techniques Used	Privacy Threat
Neuro-medical use	– Detection of brain tumor, Seizure disorder, Sleep disorder and brain swelling [16] – Control of prosthetic limbs and wheelchairs [45] – Identify Parkinson's disease and motor imaginary [57]	– Modified Wavelet-Independent Component Analysis (MwICA), Multi-Layer Feed Forward(MLFF) Neural network known as Back Propagation Network (BPN)	– Compromised system can leak patient private data to the malicious party.
Silent Speech and Auditory BCI	– Recognize unspoken speech with two words Yes and No [72] – Reconstruct silent speech through investigation of neuromuscular signals from facial and neck area [36] – Reproduce imagined speech and mouthed non-audible speech [70] – Identify the behavioral performance [55].	– Four classifiers (Support Vector Machine, Discriminant Analysis, Self-Organizing Map and Feed Forward Back-propagation) and Ensemble Network – Mel-frequency cepstral coefficient based representations, Discrete Cosine Transform – Mel Frequency Cepstral Coefficients (MFCCs), log variance Auto Regressive (AR) coefficients, Support Vector Machine (SVM), Hidden Markov Models (HMM) and k-nn classifier – Alpha and theta spectral power density curves	– Unspoken PIN, Password, Bank Information can be revealed from neuromuscular signal.
Speech Recognition	– Recognition of the first five words from the international table of the phonetic alphabet [66]	– Hidden Markov Model, Double-Tree Complex Wavelet Transform, Linear Discriminant Analysis	– Stolen neural pattern can be exploited to formulate attack to systems that requires speech verification using brainwaves.
Neural Authentication and Identification	– Identifying neural pattern and neural passwords for authentication and user identify and tracking purposes [35]	– Cosine Similarity Metric – Bilinear Transform	– Attacker can authenticate the system using the stolen neural signals which required neural authentication and can get private information from the system. – Attacker can identify person by neural patterns and can track him for malicious purpose.
Gaming and Entertainment	– Detection of amplitude peak in the EEG signal after showing stimuli (videos, pictures, alphanumeric characters etc.) on the gaming screen [59]	–Boosting algorithm for logistic regression, BCI2000 P300 classifier known as stepwise linear discriminant analysis (SWLDA) – Neural Networks: Multilayer Perceptron – Fisher's Linear Discriminant Analysis (FLDA) and Principal Component Analysis (PCA) – Linear Discriminant Analysis	– BCI games could be exploited to extract individuals' private information, such as 4-digit PINs, bank information, date of birth and location using EEG signals. – Attackers display specific videos, images or numbers and read their corresponding EEG signals to extract sensitive information.

6.2 Privacy Threat of Silent Speech and Speech Recognition Information

AlterEgo, a wearable interface that allows a user to silently converse with a computing device without any voice or any discernible movements can compromise users' privacy by neuromuscular signals in internal speech articulators [36]. This interface showed accuracy of digit recognition is 92% which may lead to privacy leakage when someone recalls PINs and passwords.

Salama, et al. showed detection of words "YES" and "NO" through the analyzing EEG signals. The work was evaluated on seven independent subjects and the EEG signal was measured by a single electrode located on the forehead of each subject. The results reported 57% success from online tests and 56% success from offline, post-analysis. [72].

A work conducted by Wester et al. tried to recognize first five words from international table of the phonetic alphabet. A specially designed EEG head cap with sixteen electrodes was used to measure brain signals from 21 individual subjects. The theory of this work was each individual brain has a specific pattern saved for each specific word and the pattern is recalled when a word is pronounced or thought about. With the proposed theory this work reported ML classifier's success rate of 45.5% [66].

In auditory brain-computer interface like [55] have been conducting researches to identify behavioral performance of patients with locked-in-state if they can recognize the sound direction to communicate with the world.

All these works demonstrated higher accuracy in learning untold/silent voice. If a malicious application use EEG responses or listen to EEG signal silently from a compromised system, then it can infer private sensitive information like PIN, password, bank account information.

6.3 Privacy Threat of User Identification and Tracking

Based on the neural pattern, each individual is unique and therefore, neural signals can be used for biometrics verification. Thus, many researchers have analyzed and recognized the potentiality of neural pattern for authentication and identification. Chuang et al. conducted a work to authenticate users based on brainwave signals [24]. In particular, they used single-channel EEG signals to record brainwave when a subject performed a custom task (e.g., singing, breathing or finger movement). Brainwaves were wirelessly transmitted to a computer application which collected and processed neural data. Their authentication system analysed similarity between brain data and training data to authenticate subjects. Their proposed authentication mechanism showed the same accuracy as multi-channel EEG authentication which was about 99% accurate. Rajagopal et al. used EEG signal to extract features as neural passwords for authentication [68]. The entire process was performed automatically without human supervision. They used an algorithm which automatically could extract neural events corresponding to individual's blinking, jaw-clenching, and eye-rolling activities. The results showed accuracy from 67% to 95% with single-trial inputs.

A compromised system can expose neural signature to adversary. Using the neural signature, attacker can impersonate thoughts of victim and use it for malicious purposes, e.g., system verification.

6.4 Privacy Threat of Gaming and Entertainment Applications

There are several brain-games available in gaming industry which are developed based on consumer-grade EEG devices. The principle of most BCI enabled games is similar to P300-speller. In such gaming, an amplitude peak in EEG signal is detected and based on that the user can interact with the gaming environment. In P300-speller, stimuli is alphanumeric character which is shown on the screen. These characters are arranged in a matrix which flashes on a screen rapidly. Users choose one character from spelled word from screen using eyes. By analyzing peaks in brain while looking at the character on screen, spelled word get identifies [59].

Kim et al. [37] measured a subject's attention and meditation level through EEG signals when a subject was playing a game. They compared the difference among all subjects' EEG signals, according to subjects' age and gender. Their results show that, in POKOPANG game, average attention level of men is lower than women, while meditation level is reversed. The concluded that women were more interested in the POKOPANG game than man. If attacker gets access to such system, he could easily classify gender of victim, and later could use this information for targeted advertisements.

Unrestricted BCI API access gave everyone capabilities to develop BCI-enabled games. A malicious party can have control over stimuli presented to users and as a consequence, attacker can learn about neural response of user and design some images, videos etc. as malicious stimuli to maximize leaked information.

7 EFFORTS ON BRAINWAVE PRIVACY ENHANCEMENT

From the studied literatures, we learnt that EEG signals obtained from consumer-grade BCI, can be used to extract private information, which can lead to serious privacy attacks. Private data extracted from brainwaves can be exploited by malicious actor to infer users' thought, memory, emotional states etc. For example, someone's memory and emotional responses might be useful for law enforcement works, criminal investigation and leaking such information will lead to devastating situation. To prevent such privacy threats, we discuss about studied efforts to enhance brainwave privacy.

7.1 Brain-Computer Interface Anonymizer

Chizeck et al. [23] introduced Brain-Computer Interface Anonymizer which can generate anonymized neural signals by filtering features to remove privacy sensitive information. It consists of two main components: the first one identifies component of recorded neural signals which is used to extract private information and quantifies the amount of information can be exposed. The second component is an analysis and validation tool which analyzes and validates the obtained information from the first component to enhance the privacy and security of BCI. Figure 5 represents the workflow of the system.

BCI Anonymizer can process brain signal components required by the application, rather than providing the whole signal packets. This implementation, resists the eavesdroppers to steal or decompose sensitive information from brain. From Section 4, we learnt about two types of attackers. BCI Anonymizer can prevent both

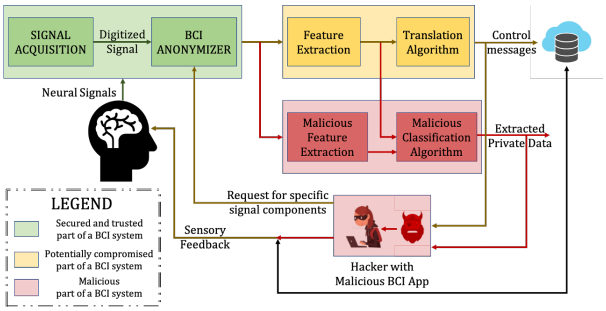


Figure 5: Brain-Computer Interface Anonymizer

types of attack scenarios by pre-processing the brain signal before it is being used or being stored in a server/database.

7.2 Privacy-Preserving Cryptographic Protocols

Agarwal et al. proposed cryptographic protocols [17] based on Secure Multiparty Computation (SMC) to perform linear regression over EEG signals from many users in a fully privacy-preserving (PP) fashion, i.e. such that each individual's EEG signals are not revealed to anyone else.

In this work, two different scenarios were considered: in the first scenario, a set of source drivers work together to train a Linear Regression (LR) model in a distributed fashion (many-party SMC). Throughout this process, none of the drivers can see the data from the other drivers in an unencrypted way at any point. At the end of the protocol, all source drivers hold encrypted shares of the trained model, and a target driver can obtain a prediction for his data by engaging in a cryptographic protocol with all of the source drivers (many-party SMC). In the second scenario, the target driver has calibration data that can be leveraged to train a personalized and more accurate model. The target driver engages in a separate cryptographic protocol with each of the source drivers (2-party SMC) to train LR models, namely as many models as there are source drivers. Each model is trained on data from one source driver, as well as on some of the calibration data from the target driver. As before, any individual's EEG data is not disclosed to anyone else. This framework allows to estimate the drowsiness of drivers as would be possible in the unencrypted case, and scales well with the number of drivers. It is the first application of commodity-based SMC to EEG data, as well as the largest documented experiment of secret sharing based SMC in general, with 15 players involved in all the computations. To do that, they presented a solution for PP training and inference with LR models in two different scenarios that are very relevant in practice, and both involve source parties and a target party.

The runtime results of this work for predicting driver drowsiness show that LR protocols and their implementation scale very nicely with an increasing number of drivers involved in the computations, and that the privately trained LR models are as accurate as those trained in the clear, i.e. without any encryption. This work shows that additive secret sharing based SMC is a viable mechanism for protecting the privacy of users in future brain-computer interface applications.

Table 6: Evaluation Summary of Possible Attack Scenarios, Attacker Type, Attacker Intuition and Potentiality and Practicality of Attacks of Currently Deployed EEG Application in Different Sectors

BCI Sectors	Private Information	Possible Attack Scenario	Attacker Type	Attacker Intuition	Potentiality of Attack
Neuro-Medical BCI Use [16, 43, 45, 53, 57, 58]	Patients' Neural Information, Pre-Processed Brain Data	Server Attack	Insider, Outsider	Learn Private Sensitive Information of Patients and use it for Advertisement purposes.	Medium
Silent Speech and Auditory BCI [36]	Processed Auditory/Hidden speech, Raw Brainwave Signal	Server Attack, Active/Passive Eavesdropping	Malicious Third-Party App	Learn Untold PIN, Password, Bank Information	High
Speech Recognition [66]	Processed/Recognized Speech, Raw Speech Data Collected From Brain	Server Attack, Passive Eavesdropping	Malicious Third-Party App	Use Stolen Speech Pattern for Critical System Verification	Medium
Neural Authentication and Identification [24, 68]	Neural Identity from Brainwave Signal	Active/Passive Eavesdropping	Malicious Third-Party App	Use Stolen Neural Signature for System Verification, Use Neural Pattern for Future Malicious Attacks	High
Gaming and Entertainment [37, 42, 59]	Keystroke Detection from Brainwave, Crack Silent Code from Brainwave	Active/Passive Eavesdropping	Malicious Third-Party App	Use Neural Pattern to Identify ATM PIN, Bank Information etc.	High

8 CONCLUSION & FUTURE RESEARCH DIRECTIONS

Literature surveyed and systematized in this paper, demonstrate that human brain can passively and actively expose sensitive private information. While a user is using a consumer-grade BCI device, being unaware of active or passive malicious activity running in background, serious privacy threats can be initiated by malicious parties, such as exposing private data like demographic information, identity of users, mental condition, emotional state, speech detection, person's interest. Later this compromised data can be used for targeted advertising on website or social networking or to run side channel attack to get bank or credit card information.

In this paper, we surveyed three active attacks [29, 56, 62] and one passive attack [81] which plays a preliminaries role to demonstrate the privacy risk of current BCI uses. All these attacks have potential success rate is presented in Table 3. Later we systematized papers which extracted private information from brainwave using different machine learning classifiers with potential accuracy which can play a role as ground truth to launch successful attacks. Table 4 demonstrates works that could potentially identify the emotional state [20, 49, 52, 63, 82], attention level [50], spoken word [25], language [75], decode English alphabets [80], left or right hand movement [51, 65], stress level [33, 34, 67], identity recognition [47] etc. At last, in Section 6, we presented BCI implementation in different sectors and the future privacy implication through a compromised system that can cause harm to the end users. Table 6 demonstrates the possible attack scenarios in daily-life BCI implementation in a compromised system and the potentiality of the attacks.

Researchers attempted to demonstrate some defense mechanisms discussed in Section 7. Brain-Computer Interface Anonymizer [23] by Chizeck et al. can generate anonymous brain-signal by eliminating private sensitive information, while Privacy-Preserving Cryptographic Protocols [17] by Agarwal et al. offers a fully privacy-preserving (PP) fashion which does not reveal signal to anyone else.

From the surveyed attack literature, neural papers with higher success rate of ML classifiers in detecting private information and defense mechanisms, we present the requirement of additional security layer in consumer-grade BCI implementation. In future, more privacy-preserving techniques need to be introduced, such as, to prevent attacks from outside attackers, manufacturers can include device license key which will allow the device to interact with the specific computing system. Inside attackers, such as, malicious third-party mobile application will not have the license key, thus will not be able to receive the brainwave signals which prevents them to exploit our brainwaves. This paper gives proper direction to security researchers to consider the privacy risks before consumer-grade EEG headsets get accessible to end users like smartwatch/smartphone.

The systematized presentation in this survey gives a clear research direction to enhance privacy in the deployment of BCI devices. While the human brain and AI will make future technology more promising, but it can lead to serious harm to the human being. Thus, to protect people from such threats, while welcoming the advanced future technology, more research effort is required to learn how much private data is possible to extract from different types of BCI devices, and how the uses of private data can be limited by adding an extra layer of filtering. Therefore, further research efforts in designing powerful attacks with higher accuracy are required to build potential defense mechanisms to counteract future privacy attacks.

ACKNOWLEDGEMENT

This research is partially supported by the National Science Foundation (NSF) under the grants: CNS-1714807, CNS-2030501, CNS-2139358.

REFERENCES

[1] 2020. Anatomy of Brain. https://www.hopkinsmedicine.org/health/conditions-and-diseases/anatomy-of-the-brain. [Accessed: 11/19/2021].
[2] 2020. B-AlertX-Series. https://www.advancedbrainmonitoring.com/products/b-alert-x-series. [Accessed: 11/19/2021].
[3] 2020. BCI in video games. https://hackernoon.com/neural-tech-and-brain-computer-interfaces-bci-in-video-games-an-overview-w1q3uge. [Accessed: 11/19/2021].
[4] 2020. Cerebellum Wiki. https://en.wikipedia.org/wiki/Cerebellum. [Accessed: 11/19/2021].
[5] 2020. Forebrain Wiki. https://en.wikipedia.org/wiki/Forebrain. [Accessed: 11/19/2021].
[6] 2020. Frontal Lobe Wiki. https://en.wikipedia.org/wiki/Frontal_lobe. [Accessed: 11/19/2021].
[7] 2020. Midbrain Wiki. https://en.wikipedia.org/wiki/Midbrain. [Accessed: 11/19/2021].
[8] 2020. Neuralink. https://neuralink.com/. [Accessed: 11/19/2021].
[9] 2020. Neurosky Mingwave Mobile EEG Headsets. https://en.wikipedia.org/wiki/NeuroSkyf. [Accessed: 11/19/2021].

[10] 2020. Nextmind visual EEG Headsets. https://www.next-mind.com/technology. [Accessed: 11/19/2021].

[11] 2020. Occipital Lobe Wiki. https://en.wikipedia.org/wiki/Occipital_lobe. [Accessed: 11/19/2021].

[12] 2020. Parietal Lobe Wiki. https://en.wikipedia.org/wiki/Parietal_lobe. [Accessed: 11/19/2021].

[13] 2020. Scientists are using brain-computer connections to restore a lost sense of touch. https://www.zdnet.com/article/scientists-are-using-brain-computer-connections-to-restore-a-lost-sense-of-touch/. [Accessed: 11/19/2021].

[14] 2020. Temporal Lobe Wiki. https://en.wikipedia.org/wiki/Temporal_lobe. [Accessed: 11/19/2021].

[15] 2020. Valve psychologist explores controlling games directly with your brain. https://venturebeat.com/2019/03/24/valve-psychologist-explores-controlling-games-directly-with-your-brain/. [Accessed: 11/19/2021].

[16] Sarah N Abdulkader, Ayman Atia, and Mostafa-Sami M Mostafa. 2015. Brain computer interfacing: Applications and challenges. *Egyptian Informatics Journal* 2 (2015), 213–230.

[17] Anisha Agarwal, Rafael Dowsley, Nicholas D McKinney, Dongrui Wu, Chin-Teng Lin, Martine De Cock, and Anderson CA Nascimento. 2019. Protecting privacy of users in brain-computer interface applications. , 1546–1555 pages.

[18] Hafeez Ullah Amin, Wajid Mumtaz, Ahmad Rauf Subhani, Mohamad Naufal Mohamad Saad, and Aamir Saeed Malik. 2017. Classification of EEG signals based on pattern recognition approach. *Frontiers in computational neuroscience* (2017), 103.

[19] Jordan J Bird, Luis J Manso, Eduardo P Ribeiro, Aniko Ekart, and Diego R Faria. 2018. A study on mental state classification using eeg-based brain-machine interface. In *2018 Int'l Conf. on Intelligent Systems (IS)*. IEEE, 795–800.

[20] A Block, W Von Bloh, and HJ1068470 Schellnhuber. 1990. Efficient box-counting determination of generalized fractal dimensions. *Physical Review A* 4 (1990), 1869.

[21] Tamara Bonaci, Ryan Calo, and Howard Jay Chizeck. 2014. App stores for the brain: Privacy & security in Brain-Computer Interfaces. In *Proceedings of the IEEE 2014 Int'l Symp. on Ethics in Engineering, Science, and Technology*. IEEE Press, 47.

[22] Stephen Chen. 2019. Forget the Facebook leak': China is mining data directly from workers brains on an industrial scale. https://www.scmp.com/news/china/society/article/2143899/forget-facebook-leak-china-mining-data-directly-workers-brains. [Accessed: 11/19/2021].

[23] Howard Jay Chizeck and Tamara Bonaci. 2014. Brain-Computer Interface Anonymizer. US Patent App. 14/174,818.

[24] John Chuang, Hamilton Nguyen, Charles Wang, and Benjamin Johnson. 2013. I think, therefore i am: Usability and security of authentication using brainwaves. In *Int'l Conf. on Financial Cryptography and Data Security*. Springer, 1–16.

[25] João M Correia, Bernadette Jansma, Lars Hausfeld, Sanne Kikkert, and Milene Bonte. 2015. EEG decoding of spoken words in bilingual listeners: from words to language invariant semantic-conceptual representations. *Frontiers in psychology* (2015), 71.

[26] Thomas Cover and Peter Hart. 1967. Nearest neighbor pattern classification. *IEEE transactions on information theory* 1 (1967), 21–27.

[27] Ruo-Nan Duan, Jia-Yi Zhu, and Bao-Liang Lu. 2013. Differential entropy feature for EEG-based emotion classification. In *2013 6th Int'l IEEE/EMBS Conf. on Neural Engineering (NER)*. IEEE, 81–84.

[28] EMOTIV. 2020. Neurotech for the Global Community. https://www.emotiv.com/. [Accessed: 11/19/2021].

[29] Mario Frank, Tiffany Hwu, Sakshi Jain, Robert T Knight, Ivan Martinovic, Prateek Mittal, Daniele Perito, Ivo Sluganovic, and Dawn Song. 2017. Using EEG-based BCI devices to subliminally probe for private information. In *Proceedings of the 2017 on Workshop on Privacy in the Electronic Society*. 133–136.

[30] Christian Herff, Dominic Heger, Adriana De Pesters, Dominic Telaar, Peter Brunner, Gerwin Schalk, and Tanja Schultz. 2015. Brain-to-text: decoding spoken phrases from phone representations in the brain. *Frontiers in neuroscience* (2015), 217.

[31] David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. 2013. *Applied logistic regression*. John Wiley & Sons.

[32] Kristin Houser. 2019. This AI Decodes Your Brainwaves and Draws What You're Looking at. https://futurism.com/the-byte/ai-draws-decodes-brainwaves. [Accessed: 11/19/2021].

[33] Houtan Jebelli, Sungjoo Hwang, and SangHyun Lee. 2018. EEG-based workers' stress recognition at construction sites. *Automation in Construction* (2018), 315–324.

[34] Seung Yeul Ji, Se Yeon Kang, and Han Jong Jun. 2020. Deep-Learning-Based Stress-Ratio Prediction Model Using Virtual Reality with Electroencephalography Data. *Sustainability* 17 (2020), 6716.

[35] Benjamin Johnson, Thomas Maillart, and John Chuang. 2014. My thoughts are not your thoughts. In *Proceedings of the 2014 ACM Int'l Joint Conf. on Pervasive and Ubiquitous Computing: Adjunct Publication*. 1329–1338.

[36] Arnav Kapur, Shreyas Kapur, and Pattie Maes. 2018. Alterego: A personalized wearable silent speech interface. In *23rd Int'l Conf. on Intelligent User Interfaces*. ACM, 43–53.

[37] Jung-Yoon Kim and W Lee. 2013. EEG signal feature analysis of smart phone game user. *Advanced Science and Technology Letters* (2013), 14–19.

[38] Nataliya Kosmyna, Alexandra Gross, and Pattie Maes. 2020. " The thinking cap 2.0" preliminary study on fostering growth mindset of children by means of electroencephalography and perceived magic using artifacts from fictional sci-fi universes. In *Proceedings of the Interaction Design and Children Conf.* 458–469.

[39] Nataliya Kosmyna, Utkarsh Sarawgi, and Pattie Maes. 2018. AttentivU: Evaluating the feasibility of biofeedback glasses to monitor and improve attention. In *Proceedings of the 2018 ACM Int'l Joint Conf. and 2018 Int'l Symp. on Pervasive and Ubiquitous Computing and Wearable Computers*. 999–1005.

[40] Gautam Krishna, Yan Han, Co Tran, Mason Carnahan, and Ahmed H Tewfik. 2019. State-of-the-art speech recognition using eeg and towards decoding of speech spectrum from eeg. *arXiv preprint arXiv:1908.05743* (2019).

[41] Ofir Landau, Aviad Cohen, Shirley Gordon, and Nir Nissim. 2020. Mind your privacy: Privacy leakage through BCI applications using machine learning methods. *Knowledge-Based Systems* (2020), 105932.

[42] Erik Andreas Larsen. 2011. *Classification of EEG signals in a brain-computer interface system*. Master's thesis. Institutt for datateknikk og informasjonsvitenskap.

[43] Mikhail A Lebedev and Miguel AL Nicolelis. 2017. Brain-machine interfaces: From basic science to neuroprostheses and neurorehabilitation. *Physiological reviews* (2017).

[44] Mu Li and Bao-Liang Lu. 2009. Emotion classification based on gamma-band EEG. In *2009 Annual Int'l Conf. of the IEEE Engineering in medicine and biology society*. IEEE, 1223–1226.

[45] QianQian Li, Ding Ding, and Mauro Conti. 2015. Brain-computer interface applications: Security and privacy challenges. In *2015 IEEE Conf. on communications and network security (CNS)*. IEEE, 663–666.

[46] Ting-Mei Li, Han-Chieh Chao, and Jianming Zhang. 2019. Emotion classification based on brain wave: a survey. *Human-centric Computing and Information Sciences* 1 (2019), 42.

[47] Wei Liang, Liang Cheng, and Mingdong Tang. 2016. Identity recognition using biological electroencephalogram sensors. *Journal of Sensors* (2016).

[48] Ke Liao, Ran Xiao, Jania Gonzalez, and Lei Ding. 2014. Decoding individual finger movements from one hand using human EEG signals. *PloS one* 1 (2014), e85192.

[49] Ningjie Liu, Yuchun Fang, Ling Li, Limin Hou, Fenglei Yang, and Yike Guo. 2018. Multiple feature fusion for automatic emotion recognition using EEG signals. In *2018 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 896–900.

[50] Ning-Han Liu, Cheng-Yu Chiang, and Hsuan-Chin Chu. 2013. Recognizing the degree of human attention using EEG signals from mobile sensors. *Sensors* 8 (2013), 10273–10286.

[51] Weifeng Liu, Xiaoming Liu, Ruomeng Dai, and Xiaoying Tang. 2017. Exploring differences between left and right hand motor imagery via spatio-temporal EEG microstate. *Computer Assisted Surgery* sup1 (2017), 258–266.

[52] Yisi Liu, Olga Sourina, and Minh Khoa Nguyen. 2010. Real-time EEG-based human emotion recognition and visualization. In *2010 Int'l Conf. on cyberworlds*. IEEE, 262–269.

[53] Joel F Lubar and Margaret N Shouse. 1976. EEG and behavioral changes in a hyperkinetic child concurrent with training of the sensorimotor rhythm (SMR). *Biofeedback and Self-regulation* 3 (1976), 293–306.

[54] Joseph G Makin, David A Moses, and Edward F Chang. 2020. Machine translation of cortical activity to text with an encoder–decoder framework. *Nature neuroscience* 4 (2020), 575–582.

[55] Alessandro Marassi, Riccardo Budai, and Luca Chittaro. 2018. A P300 auditory brain-computer interface based on mental repetition. *Biomedical Physics & Engineering Express* 3 (2018), 035040.

[56] Ivan Martinovic, Doug Davies, Mario Frank, Daniele Perito, Tomas Ros, and Dawn Song. 2012. On the feasibility of side-channel attacks with brain-computer interfaces. In *Presented as part of the 21st {USENIX} security Symp. ({USENIX} Security 12)*. 143–158.

[57] Dennis J McFarland, Janis Daly, Chadwick Boulay, and Muhammad A Parvaz. 2017. Therapeutic applications of BCI technologies. *Brain-Computer Interfaces* 1-2 (2017), 37–52.

[58] J-A Micoulaud-Franchi, Aileen Mcgonigal, Regis Lopez, Christophe Daudet, Iliana Kotwas, and Fabrice Bartolomei. 2015. Electroencephalographic neurofeedback: Level of evidence in mental and brain disorders and suggestions for good clinical practice. *Neurophysiologie Clinique/Clinical Neurophysiology* 6 (2015), 423–433.

[59] Christian Mühl, H Gürkök, D Plass-Oude Bos, ME Thurlings, L Scherffig, M Duvinage, AA Elbakyan, S Kang, M Poel, and D Heylen. 2009. Bacteria Hunt: A multimodal, multiparadigm BCI game. In *Fifth Int'l Summer Workshop on Multimodal Interfaces*. 41–62.

[60] MNRYS Murugappan, Ramachandran Nagarajan, and Sazali Yaacob. 2009. Comparison of different wavelet features from EEG signals for classifying human emotions. In *2009 IEEE Symp. on industrial electronics & applications*. IEEE, 836–841.

[61] Murugappn Murugappan, M Rizon, R Nagarajan, S Yaacob, I Zunaidi, and D Hazry. 2007. EEG feature extraction for classifying emotions using FCM and FKM. *Int'l journal of Computers and Communications* 2 (2007), 21–25.

[62] Ajaya Neupane, Kiavash Satvat, Mahshid Hosseini, and Nitesh Saxena. 2019. Brain Hemorrhage: When Brainwaves Leak Sensitive Medical Conditions and Personal Information. In *2019 17th Int'l Conf. on Privacy, Security and Trust (PST)*. IEEE, 1–10.

[63] Dan Nie, Xiao-Wei Wang, Li-Chen Shi, and Bao-Liang Lu. 2011. EEG-based emotion recognition during watching movies. In *2011 5th Int'l IEEE/EMBS Conf. on Neural Engineering*. IEEE, 667–670.

[64] Panagiotis C Petrantonakis and Leontios J Hadjileontiadis. 2009. Emotion recognition from EEG using higher order crossings. *IEEE Transactions on information Technology in Biomedicine* 2 (2009), 186–197.

[65] Gert Pfurtscheller, Ch Neuper, Doris Flotzinger, and Martin Pregenzer. 1997. EEG-based discrimination between imagination of right and left hand movement. *Electroencephalography and clinical Neurophysiology* 6 (1997), 642–651.

[66] Anne Porbadnigk, Marek Wester, and Tanja Schultz Jan-p Calliess. 2009. EEG-based speech recognition impact of temporal effects. (2009).

[67] Prima Dewi Purnamasari and Alya Fernandya. 2019. Real Time EEG-based Stress Detection and Meditation Application with K-Nearest Neighbor. In *2019 IEEE R10 Humanitarian Technology Conf. (R10-HTC)(47129)*. IEEE, 49–54.

[68] Abhejit Rajagopal, Anthony C Nguyen, and Dennis M Briggs. 2013. Neuropass: A secure neural password based on EEG. *Biomedical Engineering, 2013* (2013).

[69] Rajesh PN Rao and Reinhold Scherer. 2010. Brain-computer interfacing [in the spotlight]. *IEEE Signal Processing Magazine* 4 (2010), 152–150.

[70] Anaum Riaz, Sana Akhtar, Shanza Iftikhar, Amir Ali Khan, and Ahmad Salman. 2014. Inter comparison of classification techniques for vowel speech imagery using EEG sensors. In *The 2014 2nd Int'l Conf. on Systems and Informatics (ICSAI 2014)*. IEEE, 712–717.

[71] Sanay Muhammad Umar Saeed, Syed Muhammad Anwar, Humaira Khalid, Muhammad Majid, and Ulas Bagci. 2020. EEG based classification of long-term stress using psychological labeling. *Sensors* 7 (2020), 1886.

[72] May Salama, Loaay ElSherif, Haytham Lashin, and Tarek Gamal. 2014. Recognition of unspoken words using electrode electroencephalograhic signals. In *The Sixth Int'l Conf. on Advanced Cognitive Technologies and Applications*. 51–55.

[73] Stefan Schneegass, Youssef Oualil, and Andreas Bulling. 2016. SkullConduct: Biometric user identification on eyewear computers using bone conduction through the skull. In *Proceedings of the 2016 CHI Conf. on Human Factors in Computing Systems*. ACM, 1379–1384.

[74] M Sharanreddy and PK Kulkarni. 2013. Automated EEG signal analysis for identification of epilepsy seizures and brain tumour. *Journal of medical engineering & technology* 8 (2013), 511–519.

[75] Akshara Soman, CR Madhavan, Kinsuk Sarkar, and Sriram Ganapathy. 2019. An EEG study on the brain representations in language learning. *Biomedical Physics & Engineering Express* 2 (2019), 025041.

[76] Kazuhiko Takahashi et al. 2004. Remarks on emotion recognition from bio-potential signals. In *2nd Int'l Conf. on Autonomous Robots and Agents*. Citeseer, 186–191.

[77] V Vanitha and P Krishnan. 2016. Real time stress detection system based on EEG signals. (2016).

[78] Vladimir Vapnik. 2013. *The nature of statistical learning theory*. Springer science & business media.

[79] Xiao-Wei Wang, Dan Nie, and Bao-Liang Lu. 2014. Emotional state classification from EEG data using machine learning approach. *Neurocomputing* (2014), 94–106.

[80] YiYan Wang, Pingxiao Wang, and Yuguo Yu. 2018. Decoding english alphabet letters using EEG phase information. *Frontiers in neuroscience* (2018), 62.

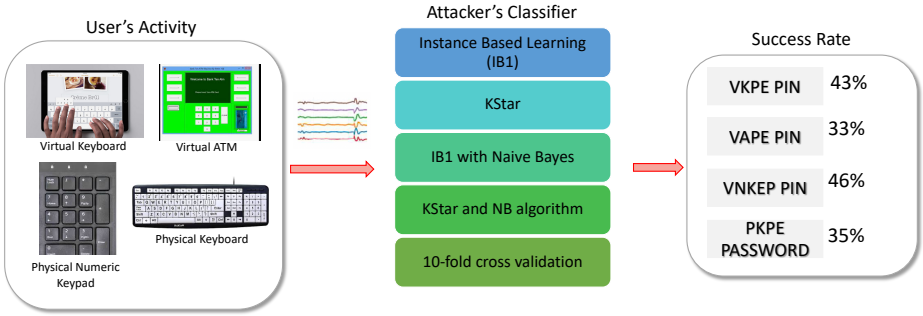
[81] Zhong Yunfan. 2019. Chinese primary school halts trial of device that monitors pupils. https://www.theguardian.com/world/2019/nov/01/chinese-primary-school-halts-trial-of-device-that-monitors-pupils-brainwaves. [Accessed: 11/19/2021].

[82] Wei-Long Zheng and Bao-Liang Lu. 2015. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Transactions on Autonomous Mental Development* 3 (2015), 162–175.

[83] Wei-Long Zheng, Jia-Yi Zhu, and Bao-Liang Lu. 2017. Identifying stable patterns over time for emotion recognition from EEG. *IEEE Transactions on Affective Computing* 3 (2017), 417–429.

APPENDIX

A STUDIED PASSIVE ATTACK



B STUDIED ACTIVE ATTACKS

