E-BIAS: A Pervasive EEG-Based Identification and Authentication System

Javad Sohankar, Koosha Sadeghi, Ayan Banerjee, Sandeep K.S. Gupta
iMPACT Lab, Arizona State University
Tempe, AZ, USA
{j.sohankar, ssadegh4, abanerj3, sandeep.gupta}@asu.edu

ABSTRACT

Security systems using brain signals or Electroencephalography (EEG), is an emerging field of research. Brain signal characteristics such as chaotic nature and uniqueness, make it an appropriate information source to be used in security systems. In this paper, E-BIAS, a pervasive EEG-based security system with both identification and authentication functionalities is developed. The main challenges are: 1) accuracy, 2) timeliness, 3) energy efficiency, 4) usability, and 5) robustness. Therefore, we apply machine learning algorithms with low training times, multi-tier distributed computing architecture, and commercial single channel dry electrode wireless EEG headsets to respectively overcome the first four challenges. With only two minutes of training time and a simple rest task, the authentication and identification performance reaches 95% and 80%, respectively on 10 subjects. We finally test the robustness of our EEG-based seamless security system against three types of attacks: a) brain impersonation, b) database hacking, and c) communication snooping and discuss the system configurations which can avoid data leakage.

Categories and Subject Descriptors

K.6 [**Management of Computing and Information Systems**]: Security and Protection—*Authentication*

General Terms

Cyber-physical systems

Keywords

Pervasive security systems, electroencephalogram

1. INTRODUCTION

Brain sensing and associated cognitive applications are fast becoming pervasive in nature due to the advent of wireless low cost easy-to-wear brain sensors that connect to mobile phones [8, 24]. This enables seamless access to a person's brainwaves which contains information that is unique to a person, nearly impossible to impersonate without invading personal space, and chaotic over time. This is markedly different from biometrics such as fingerprints, voice, and face, which can be captured without the subject's knowledge or purposefully altered [13]. Seamless availability of EEG data opens up potential usage in securing personal information in scenarios where a password may not be entered, spoken out, or remembered. For example, the notion of "hands-free" security can be imagined, when the person is driving or pre-occupied with other tasks and cannot focus on targeted security related tasks [5]. EEG-based security systems satisfy the following requirements that favor the mentioned purpose: a) universality, we always have our brain and thoughts with ourselves and hence enables pervasive security, b) uniqueness, brain signals are unique and differ from person to person potentially enabling high authentication or identification accuracy, c) permanency, some brainwave features show stable underlying behavior through time, which can be classified using machine learning techniques, but are difficult to regenerate without prior access to brain data [19], d) collectability, they can be captured by wearable sensors, and e) robustness, it's hard to hack a system through replication of brain data [3, 36]. In this paper, we propose a seamless pervasive EEG-based security system using commercially available brain sensors intended to provide authentication and identification in single user smartphones and small scale multi-user computing systems. The E-BIAS is distinguished by relatively lower training time than existing techniques and with simple mental task for the user.

There are several works in EEG-based security systems field (Section 4), which can be divided into two main categories. The first group study EEG-based security systems and evaluate their work with large number of subjects and under different conditions with comprehensive test results, but with some drawbacks that eliminate the practicality of their methods: a) usage of medical devices with tough and tedious setup procedure, b) complicated scenario tasks with long training time, c) using desktop platform and lack of mobility [9, 14, 20, 27]. On the other hand, the second group attempt to develop practical and pervasive security systems using commercial EEG sensors, but with limited performance analysis and system evaluation [13, 18]. In this work, we propose a pervasive EEG-based security system using commercial EEG headsets (Neurosky [1]) and perform thorough system evaluation based on accuracy, latency, power consumption, usability, and also robustness against three different attack types. *The main advantage of our system*

Q2SWinet'15, November 2–6, 2015, Cancun, Mexico.

DOI: http://dx.doi.org/10.1145/2815317.2815341.

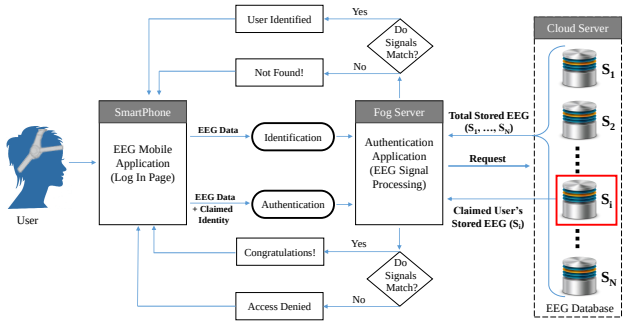

Figure 1: Model of the EEG driven security system.

is high authentication/identification rate using commercially available devices with lower training time (2 minutes) than previous works [9, 22], and implementation for pervasive usages.

1.1 System Requirements and Challenges

Previous works in pervasive interactive applications [10], brain monitoring [31, 35], and biometric systems [3, 17] address a number of system requirements that can be used to evaluate a pervasive EEG-based security system:

1) Accuracy: The chaotic nature of brain signals and impact of varying emotional states (anxiety, stress, anger, etc.) or drugs on the EEG signals can affect the success rate of security system over time. The present work uses baseline EEG signals, when a person is in rest state which is shown by recent works [3, 19] to remain stable over long periods of time. In this state, it is difficult to extract unique features of a person especially using commercial sensors. There are two possible solutions: 1) applying complex preprocessing, feature extraction, and classification methods, which might increase power consumption, and response time, and 2) collecting large training data set that can dramatically increase response time. In our research, we employ the first method. We use Fast Fourier Transform (FFT) for feature extraction and machine learning method called Naïve Bayes Classifier (NBC) that leads to high accuracy performance.

2) Timeliness: Processing EEG data using complex machine learning techniques is a time-consuming procedure especially on mobile platforms. Based on the experimental results, our system runtime on mobile phone is approximately 100 times slower than running on a desktop system. Therefore to avoid unbearable latencies as well as fulfilling real-time requirements, we use a "fog sever" based system architecture, that enables the smartphone to offload complex data processing for faster execution time [26, 35].

3) Energy Efficiency: Mobile platforms provide limited amount of resources such as energy, bandwidth, and storage capacity. The complex computation required for our system, drains the smartphone battery. Here again, using fog server (i.e., laptop) in the system architecture saves smartphone energy, and help the system handle computational and storage requirements for the application.

4) Usability: It is the ease of use of system while not deteriorating its performance. Tedious training procedures reduces system usability. Various tasks with roots in psychology and neuroscience have been designed and exploited in different works. For instance, resting/relaxing, imagining moving body parts, auditory/visual stimulation (e.g. tones, songs, colors, or images), performing mathematical operations in mind, thinking about a specific concept [9, 20], or even without doing any tasks [13]. In our research, to keep the scenario simple and acceptable by the user, we use brain signals while the user is in physically rest state. We also use a light-weight commercially available wireless EEG sensor with a single dry electrode that records signals from forehead. At last, mobile platform is used to implement sensor interface in pervasive contexts.

5) Robustness: The system should maintain required levels of security under various attacks. We evaluate our system against three types of attacks: a) impersonation, where a person attempts to imitate another person's EEG signals, b) database hacking, where the unique brain signature of a person is stolen, and c) communication snooping, where the brain features transmitted from a user over network are stolen. We use ten subjects to test the system that is comparable with recent research.

1.2 Summary of Contributions

In summary, we make the following contributions:
- We use commercially available EEG headset in our pilot system and reach higher authentication and identification accuracy with shorter training time compared to previous studies that use same type of sensors [9].
- Most researches focus on authentication or identification separately. Our system can be used to both authenticate and identify the person. Further, identification performance in our system is significantly improved (by about four times) as compared to previous works that use commercial sensors [9] over a comparable set of test subjects.
- We present an implementation of our pervasive EEG-based security system for authentication and identification in single user mobile and multiuser desktop systems, respectively and evaluate their accuracy, latency, power consumption, usability, and robustness.

1.3 Overview of Results

In our system, we use single channel commercially available EEG headset, Fast Fourier Transform (FFT) for feature extraction, naïve Bayes classifier (NBC) for classification, two minutes training time, and 10 test subjects. The authentication accuracy ranges between 81-95% depending on the length of testing samples (5-60 seconds of EEG samples), and the maximum identification accuracy reaches 80% with 50 seconds EEG test samples. After data collection the user has to wait for 800 ms to be authenticated by the system. The total energy consumption in smartphone is 1.07 J.

2. THE PROPOSED APPROACH

In this section, we discuss EEG signals, system model and usage scenarios of the system.

2.1 EEG Signals

EEG signals are electrical flows through neurons caused by brain activities which produce potential differences in order of microvolts (5-100 μV). EEG signals are captured by placing EEG electrodes on the surface of scalp [32, 34]. EEG signals are usually decomposed in several frequency bands. Each band contains signals associated with particular brain activity [9]: 0.5-3.5 Hz (δ, sleep state), 4-7 Hz (θ, drowsy state), 8-13 Hz (α, relaxation or rest state), 14-30 Hz (β, active concentration and alertness state), 30-100 Hz

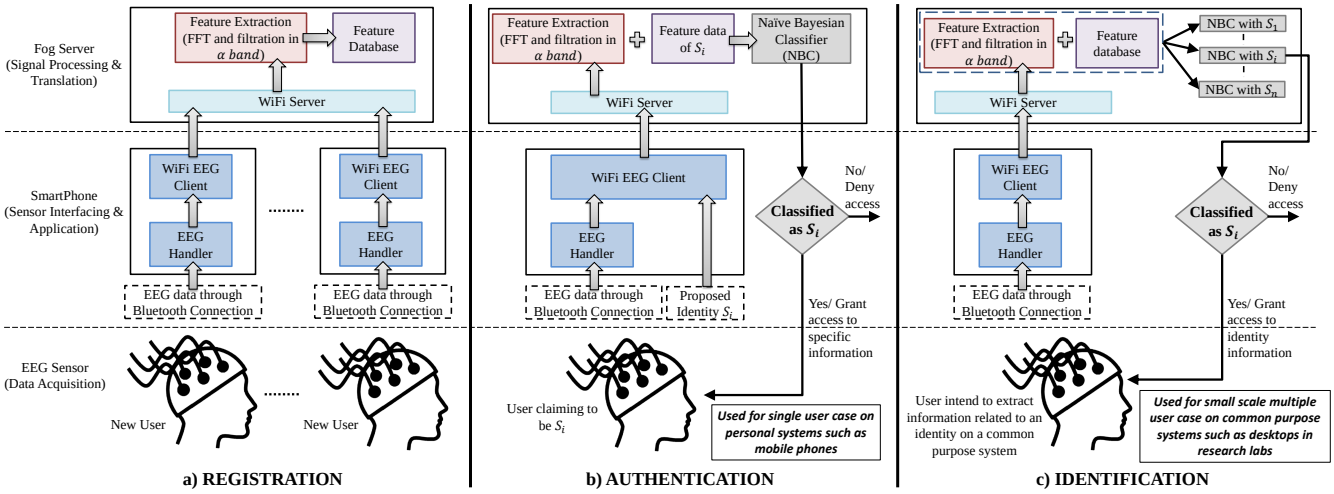

Figure 2: System architecture for different usage scenarios.

(γ, perception). In our experiments, we consider only the rest state which is marked by large variations in α wave amplitudes and it is achieved by requiring each subject to sit on a chair and relax in a distraction-free room. EEG waves are considered to be deterministically chaotic signals [23]. This means that their amplitude and duration are highly random but their unpredictability can be mimicked by non-linear dynamic learning systems such as neural networks. The significance of such chaotic nature for security system is that given a sample data set S_i and a non-linear dynamic learning system M it is extremely difficult to derive another sample data set S_i' that is accepted by the system M [6].

2.2 System Model

Figure 1 shows system model of a mobile phone security system using the proposed EEG-based solution. In this model, a mobile phone collects brain sensor data, and sends it to a fog server for extraction of EEG features and classification. The fog server uses a cloud database for the purpose of classification. The cloud database stores EEG signatures/features from each of potential users that can log into the system. The application can run in either the identification or the authentication modes which are defined below.

Authentication problem: Considering a pair of signal and identity, system should specify whether the signal matches the stored signature of the identity.

Identification problem: Considering an input signal, system should specify whether the signal matches any of the stored signatures in the system.

Identification is a more complex problem than authentication due to two main reasons: a) in identification we have to search through features of multiple subjects while in authentication, our search space is restricted to only features of a given subject, and b) the identification problem has to handle cases when input features may get classified as signatures of more than one subject.

2.3 Usage Scenarios

We envision the usage of the proposed EEG-based security technique in two scenarios: a) authentication of an individual to a single user personal mobile device (smartphone), and b) identification of an individual as a registered user for a small scale multi-user computing system such as common purpose desktops in a research facility as seen in Figure 2. We will refer to the system to which the user wants to get authenticated to as the target. The first step in both scenarios, is registration procedure.

Registration: In the registration process (Figure 2(a)), a user (S_i) is required to wear the Neurosky headset and the target collects a 2 min sample of EEG data. During these two minutes, the user is required to be in *rest* state while the user is not doing any specific mental task. The 2-minute sample is then passed to a fog server, with higher computational capacity, for extracting relevant features and storing them in a database. According to our experiments, when the target is a smartphone, it typically needs an external desktop system as the fog server for fast feature extraction. The registration process is the same for a multi-user scenario and the only difference is that the fog server, stores a database of features corresponding to different users.

Authentication: Authentication process (Figure 2(b)) is intended towards a target, which only has a single user. In this scenario, the returning user who wants to gain access to the target registered to S_i, wears the brain sensor such that the target can collect a 1-minute sample of the brain signal and send it to the fog server. The fog server, now uses the feature set of S_i and compares the collected data from the returning user by applying machine learning techniques such as NBC. If the fog server can classify the returning user as S_i, then it grants access.

Identification: The identification process (Figure 2(c)) is intended towards a target, which has multiple ($\sim$10) registered users. In this scenario, the returning user again wears the headset and the target collects a 1-minute sample. It then iterates over all possible registered users and uses machine learning techniques to classify. If the target achieves a unique classification, it grants the returning user access to the identified user account. In case of multiple or inconclusive classification the target denies access.

3. TRUST AND ATTACK MODEL

In a security system, no communication link (i.e. Bluetooth and WiFi networks) or computational units (i.e. mo-

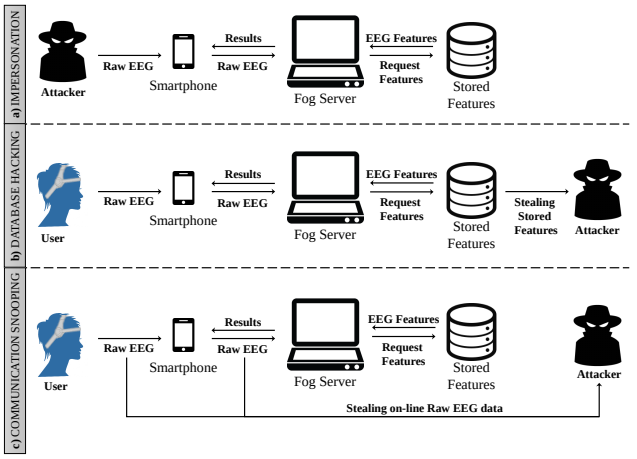

Figure 3: Three type of attacks against system.

bile devices and fog servers) is completely secure. We consider three types of attack against our system and later on propose a solution to them. In these attacks, we consider the cases where the attacker tries to fake brain signals, hack computational units and monitor communication links.

3.1 Attack Scenarios

Impersonation: As shown in Figure 3(a), impersonation occurs when the attacker wears the EEG headset and tries to mimic a user's brain signal and fool the system into granting access to her.

Database Hacking: In this type of attack, the attacker hacks the database of stored features which are the users signatures as in Figure 3(b). Then, the attacker can provide these features to the system and gain access.

Communication Snooping: The attacker monitors the communication links (e.g. between the EEG headset to smartphone or smartphone to fog server) and steals the current signal and features that user is providing to system, as seen in Figure 3(c). Later on the attacker can use these features to gain access.

3.2 Security Metrics

Here, we define some terms and metrics that we use for discussing and evaluating our security system.
a) Signature: it is the coefficients of the FFT of EEG signals in α frequency band collected during a training session.
b) input signal: it is a time series of EEG data collected during an authentication or identification session.

We define security metrics with respect to the following events in the authentication problem:
c) True Accept (TA): TA occurs when the system successfully matches the input signal from the registered user to its signature.
d) False Reject (FR): FR occurs when the system fails to match input signal from registered user to its signature.
e) True Reject (TR): TR occurs when the system correctly rejects an input signal from an unregistered user.
f) False Accept (FA): FA occurs when the system incorrectly matches an input signal from unregistered user to the signature of the registered user.

In *identification*, the definition of these events changes slightly; TA occurs when system finds the correct signature that matches the input signal of a returning registered user and FR occurs when system incorrectly rejects the signal belonging to a returning registered user. TR happens once signal from an unregistered user gets rejected and FA happens once signal from an unregistered user gets accepted as a registered user or when input signal from a returning registered user gets incorrectly matched with signature of another registered user.

The security metrics for authentication are defined in terms of rates, for example, False Reject Rate (FRR) and False Accept Rate (FAR) which are the percentage of times FR and FA occur among all trials, respectively. Finally the accuracy is defined using Equation 1:

$$Accuracy = \frac{TA + TR}{TA + TR + FA + FR} \quad (1)$$

Also, there is another metric which is mostly used in security systems known as the Half of Total Error Rate (HTER) as defined in Equation 2:

$$HTER = \frac{FRR + FAR}{2} \quad (2)$$

4. RELATED WORK

There is a wide range of research in security systems using brain signals. In this paper, we will compare E-BIAS with works that perform both authentication and identification (which are summarized in Table 2 and 3).

Authentication: One of the earliest ideas of using EEG signals for authentication dates back to Thorpe et al. research in 2005, called pass-thoughts [33]. Since then, several research exists on EEG biometric systems using medical EEG recording devices and in laboratorial conditions with satisfying results [20]. Hence, current attempts mostly focus on how to apply this method in real-life context [9]. Chuang et al. [9], at UC Berkeley have conducted a comprehensive study on the usability of commercial EEG headsets for authentication. In their experiments, they used a single channel Neurosky MindSet EEG sensor with seven different mental tasks (e.g. rest, finger, sport, song, audio, and color). They used cosine similarity method for classification. In [16], the robustness of the same authentication system (Chuang et al.) is tested against impersonation attacks. Subsequently, Ishikawa et al. [14] take the same approach (i.e. [9]), but at this time, they test the scenarios with BioSemi medical EEG headset through various electrodes arrangements (from one to sixteen channels). In [29], Riera et al. applied Fisher Discriminant Analysis to classify captured signals from two frontal electrodes in the rest state. In [18], Klonovs et al. propose an on the go authentication system on Android platform using four channels and visual stimulation (i.e. image) task. In [19], Lee et al. test the accuracy of their system using recorded signals over extended period of time (10 days - 5 months). In [4], Ashby et al. use 14-channel Emotiv headset and by using Support Vector Machine (SVM) reach nearly 100% accuracy.

Identification: Poulos et al. [28] used EEG signals in rest state from one channel and used Auto-Regressive (AR) model for feature extraction and Neural Networks (NNs) for classification to identify four subjects among 75. In [25], Paranjape et al. used Discriminant Function Analysis (DFA) on one channel EEG signals to differentiate subjects in a pool of 40 subjects. Furthermore, Hu et al. [13] suggest a ubiquitous security system using EEG signals. They use

single electrode portable EEG collection device to record signals on mobile platform and NBC is used for classification. All these systems use medical grade sensors and often require large training data sets. In [16], identification using Neurosky sensor has low accuracy of 22%.

5. SYSTEM ARCHITECTURE

In this section, we describe the architecture of a pervasive EEG-based security system where different users may attempt to access via their mobile phones. The system architecture consists of four tiers including sensors, mobile devices, fog sites, and cloud data center.

In the **first tier**, we use wireless portable EEG headset called Neurosky MindWave [1], a non-invasive and commercially available EEG headset in market, that records brain signals through a single channel from the forehead. The headset sends raw EEG signals to a target device via Bluetooth. In the **second tier**, sensor is connected to mobile phone via Bluetooth for real-time streaming of brain signals. Basic signal processing algorithms such as FFT can be implemented on available smartphones. However, they are not powerful enough to perform complex analysis on physiological data. . Near-end devices such as desktop and laptops in the **third tier** handle complex computations. These devices are called fog servers [7] and communicate with mobile devices via WiFi on same network. Fog devices can significantly reduce the workload on smartphones by handling real-time data processing, data caching, and computation offloading. The short distance between mobile devices and fog servers makes the communication fast enough with ignorable amount of latency for real-time Brain Computer Interaction (BCI) applications. However, using fog server reduces user mobility to some extent. This can be eliminated by using cloud server in the **fourth Tier**, but with higher communication latencies. The cloud data center is responsible for big data management, big data mining, multivariate machine learning, and massive parallel data processing for large scale multi-user BCI systems. In our pilot system, the cloud server is left as future extension of the system.

6. IMPLEMENTATION

A pilot system is implemented to evaluate the performance of the security system. The system platform consists of an Android smartphones and a laptop. Using smartphones enhances the mobility and ease of use of the system. A list of applied devices is seen in Table 1. At first, single dry electrode Neurosky Mindwave [1], captures the raw EEG with sampling rate of 512 Hz, and sends them to smartphone via Bluetooth. On smartphone, E-BIAS interface application receives EEG signals, sends them to a laptop (i.e. fog server), and waits for the identification/authentication results. Then, the laptop, without applying any preprocessing on the EEG data, performs FFT for feature extraction and NBC for classification. FFT is applied on each second of data to achieve α band coefficients (8-13 Hz) as the features. We decided to use α band coefficients as features for two reasons; first, α band is related to relax mode of the brain signals; and second, we tried other frequency bands and the best performance was in the α band. Finally, after classification, identification/authentication results are sent back to E-BIAS interface application on smartphone.

Table 1: List of pilot system hardware.

EEG headsets	**Neurosky MindWave mobile EEG headset** ***Electrode(s):*** One dry forehead mounted ***Sampling rate:*** 512 Hz ***Communication:*** Wireless/ Bluetooth
Mobile devices	**LG Nexus 5:** ***CPU:*** Quad-core 2.3 GHz Krait 400 ***Memory:*** 2 GB RAM
Fog servers	**Dell Latitude E6410 Laptop** ***CPU:*** Intel Core i5 CPU 2.40 Hz ***Memory:*** 4 GB RAM

For the purpose of authentication/identification, the system should be able to classify subjects based on their brain signals. At first, Adaptive Neuro Fuzzy Inference System (ANFIS) [15] was used as classifier. But, the accuracy was not satisfactory. We were only able to authenticate our subjects with 55% accuracy which is far from acceptable performance for practical security systems. So, we changed our classifier to NBC [13] which gave us higher authentication accuracy with low computational cost, and also it's results were comparable with stronger classification methods.

Bayes Theorem and conditional independence between features is the backbone of NBC and density estimation is applied on data which assigns observations to the most probable class. In general, NBC works as follows [**?**]: First, NBC estimates the densities of the predictors within each class. Then, it models posterior probabilities according to Bayes rule. For $k = 1, ..., K$ (Equation 3):

$$P(Y = k|X_1, .., X_J) = \frac{\pi(Y = k) \prod_{j=1}^{J} P(X_j|Y = k)}{\sum_{k=1}^{K} \pi(Y = k) \prod_{j=1}^{J} P(X_j|Y = k)} \quad (3)$$

where Y is the random variable corresponding to the class index of an observation, X_1 to X_J are the random predictors of an observation and $\pi(Y = k)$ is the prior probability where class index is k. Finally, NBC classifies an observation by estimating the posterior probability for each class, and assigns the observation to the class yielding the maximum posterior probability.

7. SYSTEM EVALUATION

In this section, system components are evaluated with respect to accuracy, latency, energy consumption, usability, and robustness.

Scenario description: In the experiment, we used a simple task for the sake of usability. In some works, they use scenarios which needed extensive effort from the subjects, but we just asked the subjects to be in rest state (stay calm and relax) for 2 minutes. Related works choose EEG recoding times range from 24 seconds up to 16 minutes. In our study, 2 minutes was long enough to reach desired performance and was short enough to avoid user frustration. The subjects were in silent room siting on a chair and without performing any specific mental task. We used NeuroSky mindwave sensor for capturing EEG signals from the subjects. For each subject, one session was captured. The subjects were graduate students of Arizona State University between the ages of 20 to 30 (both male and female). We had 10 subjects in our experiment which is comparable to other works. Each session data is divided into segments. We tested our system on different segment sizes from 5 to 60 seconds with 5 seconds interval.

7.1 Accuracy

Here, we discuss accuracy of the security system.

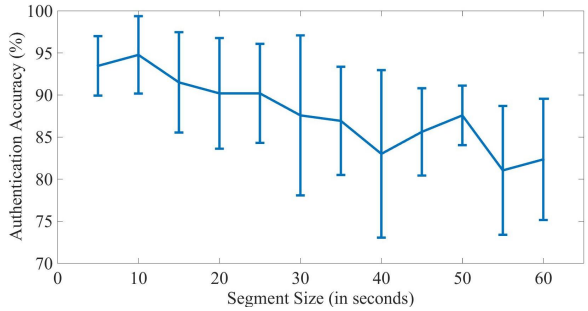

Figure 4: Average authentication rate vs. training size (i.e. segment size in seconds) over all subjects.

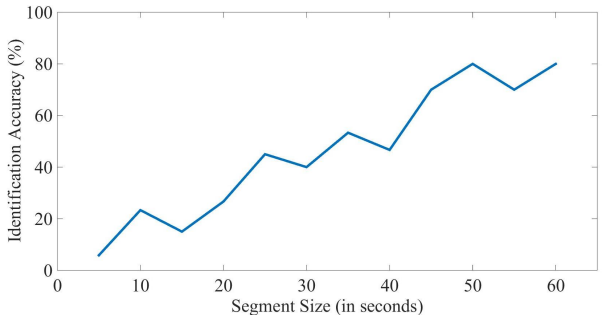

Figure 5: Identification rate vs. training size (i.e. segment size in seconds) over all subjects.

7.1.1 *Authentication*

To test the authentication mode of the system, we divide each subjects signal into segments. We choose one of these segments as the subject signature for our system database and use the rest of segments for testing the system. As explained in Section 2, for each subject, we calculate TA, FR, TR, and FA. We check each segment of a subject with all the segments of the same subject to calculate TA and FR. Afterward, we check each segment of a subject with all segments of all the other subjects for calculating TR and FA. Finally, the accuracy is calculated for different segment size varying from 5 seconds to 60 seconds with intervals of 5 seconds. In Figure 4, the average and standard deviation of the accuracy over all the subjects for different segment size is shown. We can see that the accuracy ranges from 81% at 55-second segment to 95% at 10-second segment.

7.1.2 *Identification*

Similar to the authentication mode, we divide the signals into segments and choose one of the segments as the subject signature for the system database. Then again we test all the segments from all subjects to calculate accuracy. We calculated the accuracy for different segment size varying from 5 seconds to 60 seconds with intervals of 5 seconds. In Figure 5, the average accuracy over all subjects for different segment size can be seen. The accuracy starts at 5% at 5 second segment size and increases with the segment size and reaches to 80% accuracy at 60-second segment size.

According to accuracy results shown in Figures 4 and 5, although the segment size increases, the authentication accuracy does not change a lot. But, increasing segment size leads to increase in identification accuracy, significantly. The graphs show that with segment size equal to 60s, authentication accuracy is 81% while identification reaches the highest accuracy point. Authentication is a special case of identification where the feature set just contains one sample. In this sense, identification accuracy cannot exceed authentication accuracy. So, at 60s segment size, both accuracies are around 80% where the system performance is at the highest level and remains stable from then.

7.2 Latency Analysis

The security system runs on a multi-tier platform consisting of mobile phones, laptops or desktops, and cloud data center. Here, we measure latency for each part of the system. There are two latency types for this system which are transmission latency (T) and computation latency (C). In Table 4, the latencies for each part of the system is listed.

Table 4: Latencies for each component.

EEG headset to client phone/T*	$\sim$ 1 ms
EEG app on client phone/C*	$\sim$ 1002 ms
Client phone to fog server/T	$\sim$ 425 ms
FFT and NBC on fog server (MATLAB)/C	6 ms
FFT and NBC on smartphone (Java Android)/C	840 ms
Fog server to client phone/T	$\sim$ 425 ms

* T for Transmission and C for Computation tasks.

7.3 Energy Consumption

One of the key concerns in designing any practical, efficient, and sustainable system is to keep the amount of power consumption as low as possible. Especially on mobile platforms power issue is more critical. Here, we report the power consumption on each part of the system as seen in Table 5.

Table 5: Power traces for each component.

EEG headset	50 mW
Bluetooth communication (Class 2)	2.5 mW
EEG app on client phone	345 mW
FFT and NBC on fog server (MATLAB)	$\sim$ 58 mW
WiFi communication	$\sim$ 800 mW

7.4 Usability Analysis

There are several parameters involved in usability analysis of EEG-based security systems such as type of sensors, number of channels, number of subjects, experiment duration, HTER, and accuracy. To compare the usability of our system with related works, we define a metric called **usability index**. The index is determined by assigning weights to evaluation parameters. Higher weights are assigned to parameters with more important roles in usability of the system. Tables 2 and 3 list test setup parameters, results, and usability indexes for authentication and identification, respectively. All the experiments use rest task for EEG recording. In our study, we set the parameter weights (w_i) for usability index of authentication process as follows: "scenario duration": 3.0, "accuracy": 2.0, "number of channels": 2.0, "device type": 2.0, "number of subjects": 1.0, and "HTER": 0.5. In identification, number of subjects play an important role in the system performance, so higher wight is assigned to it, "scenario duration": 3.0, "accuracy": 2.0, "number of subjects": 2.0, "device type": 2.0, and "number of channels": 1.0. For usability analysis, number of subjects and accuracy

Table 2: Authentication Results.

Reference	Classifier	Channel(s)	Subject(s)	HTER	Accuracy	Scenario Duration	Device	Usability Index
[29]	FDA	2	40	10.9%	-	9-12 minutes	Medical (ENOBIO)	5.4
[14]	Cosine Similarity	1	10	-	85-90%	12 minutes	Medical (BioSemi)	5.6
[22]	Spectral Distribution Analysis	1	23	11%	79%	30 minutes	Commercial	5.7
[30]	FDA*	2	51	1.7%	-	8-16 minutes	Medical (ENOBIO)	5.8
[2]	NNs	1	10	-	70-87%	5 minutes	Medical (g.tec)	6.1
[19]	LDA*	1	4	-	87-100%	5 minutes	Medical	6.3
[12]	NNs	3	6	-	95%	100 seconds	Medical (Bio Amps)	6.5
[4]	SVM*+ voting	14	5	3%	97-100%	150 seconds	Commercial (Emotiv)	6.9
[9]	Cosine Similarity	1	15	14%	85%	12 minutes	Commercial (Neurosky)	7.3
Present work	NBC	1	10	2-9%	81-95%	2 minutes	Commercial (Neurosky)	8.5

* SVM: Support Vector Machines, LDA: Linear Discriminant Analysis, and FDA: Fisher Discriminant Analysis.

Table 3: Identification Results.

Reference	Classifier	Channel(s)	Subject(s)	Accuracy	Scenario Duration	Device	Usability Index
[9]	Cosine Similarity	1	15	22%	12 minutes	Commercial (Neurosky)	3.9
[11]	NNs	3	15	60%	200 seconds	Medical	4.3
[28]	LVQ NNs*	1	4(+75 intruders)	76-88%	3 minutes	Medical	5.0
[13]	NBC	1	11	66-100%	4 minutes	Medical (NeXus-4/Mind Media)	5.1
[21]	NNs	1	10	80-97%	24 seconds	Medical	6.1
[25]	DFA*	1	40	79-85%	68 seconds	Medical	7.2
Present work	NBC	1	10	80%	2 minutes	Commercial (Neurosky)	7.5

* LVQ NNs: Learning Vector Quantization Neural Networks, and DFA: Discriminant Function Analysis.

are directly proportional to usability index, and we define their corresponding factors as seen in Equation 4:

$$f_i = \frac{x_i}{max(X)} \tag{4}$$

where X is set of values for a specific parameter and x_i is the parameter value for study i ($x_i \in X$). On the other hand, remaining parameters (i.e, number of channels, HTER, and scenario duration) and usability index are inversely proportional, as defined in Equation 5:

$$f_i = 1 - \frac{x_i}{max(X)} \tag{5}$$

In addition, the factor value for medical and commercial device types are set to 0.0 and 1.0, respectively. To calculate usability index, each weight is multiplied to its corresponding factor value, summed with other factor products, and finally divided by sum of the weights as seen in Equation 6. Higher usability indexes indicate higher system performance. As seen in the last column of Tables 2 and 3, the present work has higher usability index compared to previous works (for more readable representation, usability indexes are multiplied by ten).

$$Usability\ Index = \frac{\sum f_i w_i}{\sum w_i} \tag{6}$$

7.5 Security Analysis and Robustness

In this section, we evaluate our system against the three attack scenarios.

Impersonation: Robustness against impersonation attacks depends on FAR. We performed a study where 10 authentication systems were setup each for a given subject. For a particular authentication system, we attempted authentication using the EEG data from the other 9 subjects. Table 6 shows the number of times a wrong person was authenticated to a system out of 24 trials. The rows in Table 6 denote authentication systems for subjects 1 to 10, while the columns indicate authentication attempts from subjects 1 - 10. We derive two important conclusions from this experiment: a) the FAR is pretty low for most of the individuals, and b) the false accept events are independent. This means that false accept is rare and it is unlikely that a person gets consecutive wrong accesses (in most of the cases there were only 1 false accept while there is no evidence of consecutive false accepts). Hence, impersonation attacks can be countered by requiring multiple consecutive successful attempts. For example, if we require that the user has to have three consecutive successful authentication to gain data access then this reduces the FAR from 0.05 to $(0.05)^3$ = 0.000125, however, latency increases only linearly.

Database Hacking: A solution to this attack is to check if the features are exactly (or extremely highly) the same as the stored features. In that case, system realizes that the provided features are stolen from the database, and it will deny access. This is based on the chaotic nature of the brain signals, where it is almost impossible that another signal has exactly the same features as the stored features of the user, although the two features can belong to the same class. From our experiments we see that features collected from the same person at different times have non-zero error with respect to the features used during training, hence supporting our claim.

Communication Snooping: A solution to this type of attack is to update the user signature in the database with the latest features that were granted access. In this situation, same as the solution in database hack attack, when the attacker provides the features to the system because its same as the stored features, it will not get access. Another solution is to keep all the feature sets which the system has accepted. In this case, as the attacker tries to get access using the stolen features, it will get checked against all the feature sets in the database and access will be denied.

8. CONCLUSIONS AND DISCUSSION

In this paper, we designed and implemented a pervasive EEG-based security system model for both identification and authentication. The model-based pilot system eliminates a number of challenges in developing large scale brain mobile interface applications such as ease-of-use, portability, and performance. We have conducted extensive experiments using commercially available wearable Neurosky brain sensors. Evaluation results reveal that the system achieves 95%

Table 6: Number of FA events out of 24 trials.

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
S_1	-	0	**1**	0	0	**1**	0	0	0	0
S_2	0	-	0	0	0	**1**	0	0	0	0
S_3	0	0	-	0	0	**1**	0	0	**1**	0
S_4	0	0	0	-	0	0	0	**1**	0	0
S_5	0	0	0	0	-	0	0	0	0	0
S_6	0	0	0	0	0	-	0	0	**1**	0
S_7	0	0	0	0	0	**2**	-	0	0	0
S_8	0	0	0	**1**	**1**	0	0	-	0	0
S_9	0	0	0	0	0	0	0	0	-	0
S_{10}	**1**	0	0	0	0	**2**	0	0	**1**	-

accuracy for authentication and 80% accuracy for identification with a relatively lower training time of 2 mins. With 10 potential users the identification scheme requires $\approx 840ms$ on top of the monitoring time of 2 mins and can be implemented as a real time application in mobile phones. Large scale implementation of this scheme in systems such as bank transactions, and border control is envisioned in the future.

9. ACKNOWLEDGMENTS

This work has been partly funded by CNS grant #1218505, IIS grant #1116385, and NIH grant #EB019202.

10. REFERENCES

[1] Neurosky body and mind quantified. neurosky.com.

[2] M. K. Abdullah, K. S. Subari, J. L. C. Loong, and N. N. Ahmad. Analysis of effective channel placement for an eeg-based biometric system. In *Biomedical Engineering and Sciences, IEEE EMBS Conference on*, pages 303–306, 2010.

[3] A. Almehmadi and K. El-Khatib. The state of the art in electroencephalogram and access control. In *Communications and Information Technology (ICCIT), 2013 Third International Conference on*, pages 49–54. IEEE, 2013.

[4] C. Ashby, A. Bhatia, F. Tenore, and J. Vogelstein. Low-cost electroencephalogram (eeg) based authentication. In *Neural Engineering, 5th International IEEE/EMBS Conference on*, pages 442–445, 2011.

[5] A. Banerjee, S. K. S. Gupta, and K. K. Venkatasubramanian. PEES: physiology-based end-to-end security for mhealth. In *Proceedings of the 4th Conference on Wireless Health*, page 2. ACM, 2013.

[6] M. Barreno, B. Nelson, R. Sears, A. D. Joseph, and J. D. Tygar. Can machine learning be secure? In *Proceedings of ACM Symposium on Information, computer and communications security*, 2006.

[7] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli. Fog computing and its role in the internet of things. In *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*, pages 13–16. ACM, 2012.

[8] A. Campbell and T. Choudhury. From smart to cognitive phones. *Pervasive Computing, IEEE*, 11(3), 2012.

[9] J. Chuang, H. Nguyen, C. Wang, and B. Johnson. I think, therefore i am: Usability and security of authentication using brainwaves. In *Financial Cryptography and Data Security*, pages 1–16. Springer, 2013.

[10] A. L. S. Ferreira, L. C. de Miranda, E. E. C. de Miranda, and S. G. Sakamoto. A survey of interactive systems based on brain-computer interfaces. *SBC Journal on Interactive Systems*, 4(1):3–13, 2013.

[11] C. Hema and A. A. Osman. Single trial analysis on eeg signatures to identify individuals. In *Signal Processing and Its Applications (CSPA), 2010 6th International Colloquium on*, pages 1–3. IEEE, 2010.

[12] C. R. Hema, M. Paulraj, and H. Kaur. Brain signatures: a modality for biometric authentication. In *Electronic Design, ICED International Conference on*, pages 1–4. IEEE, 2008.

[13] B. Hu, C. Mao, W. Campbell, P. Moore, L. Liu, and G. Zhao. A pervasive eeg-based biometric system. In *Proceedings of international workshop on Ubiquitous affective awareness and intelligent interaction*. ACM, 2011.

[14] Y. Ishikawa, C. Yoshida, M. Takata, and K. Joe. Validation of eeg personal authentication with multi-channels and multi-tasks.

[15] J.-S. Jang. Anfis: adaptive-network-based fuzzy inference system. *Systems, Man and Cybernetics, IEEE Transactions on*, 23(3):665–685, 1993.

[16] B. Johnson, T. Maillart, and J. Chuang. My thoughts are not your thoughts. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, pages 1329–38, 2014.

[17] W. Khalifa, A. Salem, M. Roushdy, and K. Revett. A survey of eeg based user authentication schemes. In *Informatics and Systems, 8th International Conference on*. IEEE, 2012.

[18] J. Klonovs, C. K. Petersen, H. Olesen, and A. Hammershoj. Id proof on the go: Development of a mobile eeg-based biometric authentication system. *Vehicular Technology Magazine, IEEE*, 8(1):81–89, 2013.

[19] H. J. Lee, H. S. Kim, and K. S. Park. A study on the reproducibility of biometric authentication based on electroencephalogram (eeg). In *Neural Engineering, 2013 6th International IEEE/EMBS Conference on*, pages 13–16, 2013.

[20] S. Marcel and J. d. R. Millán. Person authentication using brainwaves (eeg) and maximum a posteriori model adaptation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(4):743–752, 2007.

[21] G. Mohammadi, P. Shoushtari, B. Molaee Ardekani, and M. B. Shamsollahi. Person identification by using ar model for eeg signals. In *Proceeding of World Academy of Science, Engineering and Technology*, volume 11, pages 281–285, 2006.

[22] I. Nakanishi, S. Baba, and C. Miyamoto. Eeg based biometric authentication using new spectral features. In *Intelligent Signal Processing and Communication Systems. International Symposium on*, pages 651–654. IEEE, 2009.

[23] E. Niedermeyer and F. L. da Silva. *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.

[24] K. S. Oskooyee, A. Banerjee, and S. K. S. Gupta. Neuro movie theatre: A real-time internet-of-people based mobile application.

[25] R. Paranjape, J. Mahovsky, L. Benedicenti, and Z. Koles. The electroencephalogram as a biometric. In *Electrical and Computer Engineering, 2001. Canadian Conference on*, volume 2, pages 1363–1366. IEEE, 2001.

[26] M. Pore, K. Sadeghi, A. Banerjee, and S. K. S. Gupta. Enabling real-time collaborative brain-mobile interactive applications on volunteer mobile devices. In *The 2nd ACM Workshop on Hot Topics in Wireless, Paris, France*, 2015.

[27] M. Poulos, M. Rangoussi, and N. Alexandris. Neural network based person identification using eeg features. In *Acoustics, Speech, and Signal Processing. Proceedings., IEEE International Conference on*, volume 2, pages 1117–20, 1999.

[28] M. Poulos, M. Rangoussi, N. Alexandris, et al. Person identification from the eeg using nonlinear signal classification. *Methods of information in Medicine*, 41(1), 2002.

[29] A. Riera, A. Soria-Frisch, M. Caparrini, I. Cester, and G. Ruffini. 1 multimodal physiological biometrics authentication. *Biometrics: Theory, Methods, and Applications*, pages 461–482, 2009.

[30] A. Riera, A. Soria-Frisch, M. Caparrini, C. Grau, and G. Ruffini. Unobtrusive biometric system based on electroencephalogram analysis. *EURASIP Journal on Advances in Signal Processing*, 2008:18, 2008.

[31] S. Sharieh, A. Ferworn, V. Toronov, and A. Abhari. An ad-hoc network based framework for monitoring brain function. In *Proceedings of the 11th communications and networking simulation symposium*, pages 49–55. ACM, 2008.

[32] D. Tan. Brain-computer interfaces: applying our minds to human-computer interaction. In *Workshop at CHI*, 2006.

[33] J. Thorpe, P. C. van Oorschot, and A. Somayaji. Pass-thoughts: authenticating with our minds. In *Proceedings of workshop on New security paradigms*, pages 45–56. ACM, 2005.

[34] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791, 2002.

[35] J. K. Zao, T.-T. Gan, C.-K. You, et al. Pervasive brain monitoring and data sharing based on multi-tier distributed computing and linked data technology. *Frontiers in human neuroscience*, 8, 2014.

[36] A. Zúquete, B. Quintela, and J. P. S. Cunha. Biometric authentication with electroencephalograms: Evaluation of its suitability using visual evoked potentials. In *Biomedical Engineering Systems and Technologies*. Springer, 2011.