symmetry

MDPI

Article

Authentication with What You See and Remember in the Internet of Things

Wayne Chiu [1], Chunhua Su [2], Chuan-Yen Fan [3], Chien-Ming Chen [4] and Kuo-Hui Yeh [1,2,5,*]

1 Department of Information Management, National Dong Hwa University, Hualien 97401, Taiwan; 410235014@gms.ndhu.edu.tw
2 Division of Computer Science, University of Aizu, Aizu-Wakamatsu 965-8580, Fukushima Prefecture, Japan; chsu@u-aizu.ac.jp
3 Department of Information Management, National Taiwan University of Science and Technology, Taipei 10607, Taiwan; D10409101@mail.ntust.edu.tw
4 College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China; chienming.taiwan@gmail.com
5 Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung 804, Taiwan
* Correspondence: khyeh@gms.ndhu.edu.tw; Tel.: +886-3-8903117

Received: 7 August 2018; Accepted: 22 October 2018; Published: 23 October 2018

Abstract: The Internet-of-Things (IoT) is an emerging paradigm seamlessly integrating a great number of smart objects ubiquitously connected to the Internet. With the rise in interest in the IoT, industry and academia have introduced a variety of authentication technologies to deal with security challenges. Authentication in IoT involves not only shifting intelligent access control down to the end smart objects, but also user identification and verification. In this paper, we build an authentication system based on brainwave reactions to a chain of events. Brainwaves, as external signals of a functioning brain, provide a glimpse into how we think and react. However, seen another way, we could reasonably expect that a given action or event could be linked back to its corresponding brainwave reaction. Recently, commercial products in the form of wearable brainwave headsets have appeared on the market, opening up the possibility of exploiting brainwaves for various purposes and making this more feasible. In the proposed system, we use a commercially available brainwave headset to collect brainwave data from participants for use in the proposed authentication system. After the brainwave data collection process, we apply a machine learning-based approach to extract features from brainwaves to serve as authentication tokens in the system and support the authentication system itself.

Keywords: authentication; brainwave; wearable; machine learning

1. Introduction

IoT stands for the Internet of Things, a concept that has become one of the most oft-mentioned topics in the computer domain. The prospect of an IoT-based economic system has brought small, embeddable devices with Internet connectivity and data collection capabilities to the market and they are becoming increasingly commonplace in our daily life. These devices are mostly resource-limited and, in many cases, implemented with microcontrollers and equipped with little usable memory [1]. Modern society is both fast-paced and competitive, and most companies focus more on device functionality than on the underlying security framework and mechanism [2]. Such implementation leaves the door open to security breaches.

Most IoT devices lack a fully functional user interaction interface, which makes the implementation of traditional authentication schemes in the IoT impractical. Those schemes only

authenticate the user at the moment of login, which is not suitable for most IoT devices [3] as they operate continuously for a long time following authentication. The entire duration of a device's operation, from the moment of login onward, should be guaranteed and protected. Continuous authentication is the best candidate for ensuring the whole session, meaning authentication and auditing in the IoT network must be ongoing.

Some implementations of continuous authentication for the IoT and smart devices have been developed, such as Google SmartLock and Microsoft Dynamic Lock. Google SmartLock is a combination of environmental-based, behavior-based, and biometric-based continuous authentication schemes [4], while Microsoft Dynamic Lock is environmental-based [5]. In this research, we focus on the biometric-based continuous authentication model. In this early stage of biometric-based continuous authentication, the model is generally considered impractical because of the costly and non-portable apparatus required [6] and the lack of computing power. With the advancement of information technology, however, the infrastructure has become more functional and practical for continuous authentication, and chipsets are larger, more energy-efficient, and offer a better performance than before. Wearables operate with biometric sensors, and a variety of communicating interfaces are attainable. Biometric information can be easily collected and exchanged. For these reasons, continuous biometric-based authentication has become practical.

The brain, as the most sophisticated organ, serves as the command center of the human body. It might be said to be the origin of every action we take. Brainwaves are the external signals of a functioning brain, and thus provide a glimpse into what we are thinking and doing [7]. A good example is that, when an individual closes his/her eyes, we can observe a slightly shivering alpha wave embedded in the brainwaves [8]. This fact sparks our curiosity and leads us to the assumption that any action or recognition will have a distinct type of brainwave waveform. Recently, more wearable brainwave-sensing devices have entered the market, lowering the barriers to collecting brainwave data and performing research on brainwaves. As this trend continues, further such usage of brainwave data can be anticipated.

Therefore, in this paper, we propose an authentication system based on brainwave reactions to a chain of events. In the proposed system, we use a commercial and wearable brainwave headset to collect brainwave data from participants for use in the proposed authentication system. After the brainwave data collection process, we apply a machine learning-based approach to extract features from brainwaves to serve as authentication tokens in the system and support the authentication system itself. In Section 2, we introduce the state of the art of IoT authentication and discuss relevant studies. Then, we present the detailed processes of our proposed authentication system with the experimental results in Section 3; finally, we give concluding remarks in Section 4.

2. The State of the Art of IoT Authentication

IoT devices are so prevalent in our daily life and so well-embedded in the environment, but security and privacy are still issues of concern [9]. Every IoT smart device we use generates sensitive individual data and can record our every movement. Consider how the data recorded by a digital door lock could be used to draw inferences about a homeowner's lifestyle, for example. Finding a way to provide security mechanisms that are both efficient and simple enough for users is, therefore, an urgent priority.

Authentication is a series of procedures for confirming the genuineness and legitimacy of an entity (such as an individual, computer, device, machine, or sensor), comprising identification along with a strict access control mechanism for authorized and non-manipulated entities. In traditional authentication, the operation processes of entity identification typically rely on usernames and passwords, which usually are not secure. Also, a method may not work with unattended end devices. In recent years, considerable research has gone into refining authentication cryptographic mechanisms as a more robust way of securing communication over IoT networks. Investigation of the design of strong cryptography for embedded systems is on the rise in the race to protect against counterfeiting,

firmware tampering, and illegal access, as reflected by works such as [10]. Hardware-supported authentication provides secure storage for maintaining IoT devices' credentials, which may be released to the microcontroller for authentication by checking the integrity of credentials.

On the other hand, lightweight cryptography for authentication is a core technology for securing the IoT and IoT-enabled services that can be implemented in constrained environments including RFID tags, sensors, wearable healthcare devices, and so on. For IoT devices, the lightweight cryptography should achieve high efficiency for end-to-end communication and security. At the same time, it has to offer easy implementation with low-resource devices. Authenticated encryption is very useful for IoT applications because it can simultaneously provide both confidentiality and authenticity of data. The AES-GCM CAESAR algorithm uses AES in the counter (CTR) mode and a Galois mode of authentication [11], and is one of the most popular authenticated encryption methods. CAESAR is a competition for authenticated encryption, focusing on security, applicability, and robustness. Competitors aim at developing new and improved versions of AEAD schemes, to overcome the shortcomings of the AES-GCM algorithm [12]. CAESAR began in 2014 with 54 initial submissions. Among the candidates, Joltik and Deoxys are based on tweakable block ciphers and use linear and lightweight transformations that are efficient for implementation on IoT applications. They show excellent hardware and software performance and excellent security. CLOC and SILC are AES-based schemes that provide partial nonce misuse resistance and can achieve an acceptable security level. They offer outstanding performance for small size datasets with minimal pre-computation and low memory requirements, which is suitable for IoT device authentication. Furthermore, several studies [13–15] have investigated the possibility of security enhancement or efficiency improvement for key agreement and secure communication for multi-modal IoT networks.

2.1. Transparent Authentication and IoT

Traditional authentication schemes require interrupting user operations to have users enter their credentials, such as messages prompting users to input a passcode to continue. Such methods are not suitable for IoT devices, not only because of the growing demand for fluid UI (User Interface) and UX (User Experience), but more importantly because a large number of devices lack a sufficient interface for doing so. Furthermore, with the trend toward Always-On devices, any intrusion by unauthorized users must be prevented. Because of the demands for the enhancement of both security and user experience, transparent authentication becomes a better mechanism for IoT devices.

Traditional authentication usually authenticates users only at the initial authorizing session, which may make them vulnerable. Consider a user who leaves his/her computer logged in so that someone unauthorized can access it. Traditional authentication can only ensure validity at the moment of authentication; after that, it cannot guarantee whether the user is authorized or unauthorized. Transparent authentication can provide the security of the whole working session. Crawford pointed out that transparent authentication has the following benefits over traditional methods [16]:

Effortless: Since the behavioral biometrics are gathered in the background, during regular device use, the user does not need to interrupt his/her tasks to authenticate.

Fine-grained access control: Traditional authentication mechanisms allow for point-of-entry authentication; once the user has provided the correct shared secret, all data and functionality on the device are accessible. Transparent authentication has the ability to provide access control on a per-task or per-data basis.

Continuous: The behavioral biometrics may be selected to take advantage of the most frequently performed tasks such as typing or speaking. In this way, there is a rich source of information used to authenticate, which supports a continuous authentication model.

With transparent authentication, any unfamiliar event or behavior that is out of the ordinary (vis-à-vis the original user's behavior) will raise a flag with the authentication system, which will either request a second authentication or initiate a lockdown. Also, transparent authentication can improve the user experience [17]. Most IoT devices do not even have a sufficient user interface to

allow the keying in of credentials or perform manual authentication. Moreover, cases abound in which traditional authentications can fail to identify valid users, such as when a valid user forgets his/her credentials. With transparent authentication, credentials are defined and given by users' bio-features and behavior. No manual interaction, or, at least, minimal interaction, is required between a human and the authentication system. So, how do we put transparent authentication into IoT practice? Something that is in such proximity to users that it can collect data about their bio-features and behavior anytime, thereby creating a unique credential for each user based on the data, is possible. Wearables have emerged as a promising solution.

2.2. Wearables to Aid Transparent Authentication

One variant of IoT device, albeit one that is relatively small in number, is the wearable device.

The term describes all IoT devices one can wear, such as smart watches, smart bands, or smart glasses. Such products are almost the same as conventional consumer electronics, making them familiar to users. In 2016, CCS Insight [18] predicted that the year 2020 would see 411 million smart wearables being sold, including 97 million pieces of eyewear, 9 million "hearables," 164 million wristbands, 25 million wearable cameras, 110 million watches, and 4 million tokens. There is little doubt that smart wearables will become commonplace and increasingly ubiquitous in the future.

Some companies developed the use of wearables as a part of the transparent authentication on devices, such as the Google Smart Lock and Microsoft Dynamic Lock. Google Smart Lock provides a mechanism for a smartphone to detect its environment and determine whether the phone is near a valid user or in a safe place. When the phone detects a user's smart wristband, smartwatch, or Bluetooth headphone nearby, the user can use the phone directly without inputting a password [4]. The Microsoft Dynamic Lock is similar; although it does not provide an unlocking mechanism, it will detect whether a valid user is still nearby or not, to determine whether to lock the computer temporarily or not [5]. In light of the flourishing and expanding coverage of wearable devices, it behooves us to demonstrate a mechanism that provides transparent authentication capability for the IoT.

2.3. Brainwaves and Authentication

Recently, BCI (Brain-Computer Interfaces) have prospered. However, their use in the realm of authentication is still relatively novel. Tulceanu presents using emotional stimulation, whereby participants listen to words related to circumstances in which they experienced contempt, disgust, fear, worry, pleasure, affection, love, pride, hope, and sadness [19]. Then brainwave signals were collected and processed. The research Tulceanu performed is rare for using stimulation and participant reactions as authentication tokens. Other works are focusing on visual stimuli and the reaction of brainwaves [20]. Finally, one way of testing is to let participants hear a steady "beep" sound to find out the differences in their reactions [21].

Most recent research works focus on stimulation and reaction. However, how memories combine with reaction is a different story. We believe that memories, which are different between individuals and thus hard to imitate, can be the best token. Combining visual stimulation and bringing out one's memories has not been done yet. Although Tulceanu presents a method somewhat similar but using voices [19], eliciting a specific reaction from a participant is nearly impossible. However, with images, by contrast, it was possible.

2.4. Authentication Technologies for IoT Communications

To secure IoT devices, we need to apply different protection to various aspects of system robustness. Fundamental secure components, like Secure Element, ARM TrustZone, or Trust Execution Environment, will be handy when doing cryptographic operations. Hardware secure modules with optimized circuits also significantly reduce power consumption while providing better performance. Secure firmware is required to protect a system from malicious tampering and provide proper isolation between core modules and applications. However, authentication in IoT is a challenge that must be

overcome to establish a trustable IoT environment further. We, therefore, need to address the challenge regarding authentication in IoT devices.

Typically, an IoT device has no display or physical interface. Therefore, even the most primitive password solution does not seem to work conveniently. Also, it is easy for attackers to hijack IoT devices, like IP cameras, due to misconfiguration, an empty password, or reliance on a default password. Authentication is essentially the front door of a device, and we need reliable solutions that are superior to passwords to block unauthorized entities from accessing the device.

In an IoT environment (e.g., Figure 1), there are standalone devices, cloud-connected devices, cloud services, and users. The authentication between users and cloud services is not limited to the IoT field, and some standards focus on it, e.g., FIDO Universal 2nd factor (U2F) [22] and Web Authentication (W3C) [23]. A user will typically possess a primary interaction device, i.e., a smartphone or wearable device, that has a user interface and supports local authentication through various authentication models such as biometrics, PIN, or gestures.

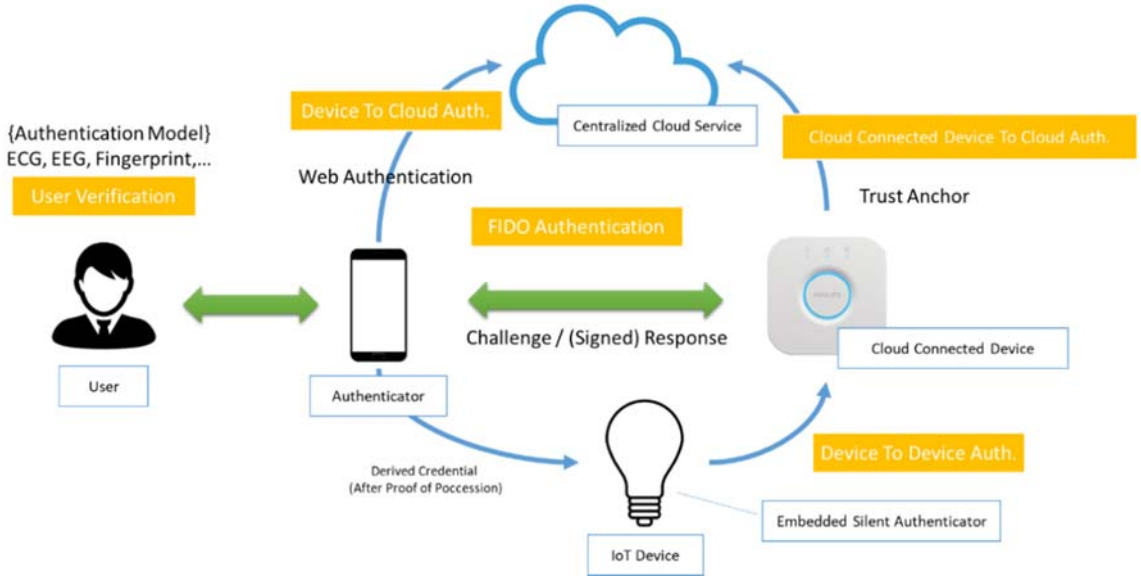

Figure 1. Emerging authentication technologies for IoT networks.

The primary interaction device acts as an authenticator that allows a user to strongly authenticate him-/herself on other IoT devices with ease. Also, crypto operations are backed by Trust Execution Environment or Secure Element, and thus have a high level of tamper-resistant protection. The trust relationship between the owners and their primary interaction devices creates a scoped authentication flow, meaning a user does not directly authenticate with the IoT devices by entering a password or PIN. Instead, the user is verified first by the primary interaction device. If the user verification is successful, the secure component will be triggered and unlock the private key to compute the signed response and proceed with the authentication flow. This procedure avoids direct interaction with any suspicious devices that could lead to accidental exposure of credentials.

Standalone devices are devices that do not have any uplink to cloud services, for example, light bulbs or smart locks. A user will sometimes need to gain access to these devices to manipulate them. As standalone devices are accessible from outside the security perimeter, they are thus likely to be exposed to compromised devices or devices with malicious software installed. Cloud-connected devices rely on cloud services and use cloud services to update configurations and transfer command sets. Also, cloud services are likely to add extra value to IoT products. Therefore, a trust relationship must be established between cloud services and cloud-connected devices to allow Device-to-Cloud authentication. As mentioned previously, there are well-implemented authentication mechanisms (such as FIDO U2F and W3C) to be taken into account between users and cloud services.

Furthermore, a user usually would like to have access to cloud-connected devices even when there is no network connection to the cloud services, or the cloud services are unreachable. Hence, the user has a demand for direct authentication to the cloud-connected device. Finally, there are devices connected to other devices, which requires Device-to-Device authentication.

As in the diagram presented in Figure 1, in an IoT environment, various entities (users, authenticators, standalone IoT devices, cloud-connected devices, cloud services) exist. Depending on the context, there will be different authentication mechanisms being deployed [24,25]:

- User verification to authenticator: A user carries the primary interaction device, such as a smartphone or wearable device. These devices provide the user verification interface and have a secure element on them to support authentication operations. The user verification is used to grant access to the legitimate user. There are various options upon user verification, for example, facial recognition, ECG, EEG, and fingerprints, all of which can also ensure that the legitimate user is actually physically present during the verification phase and hence protect the authentication from remote attack. This method is called "Local Authentication," a term that applies when the user verification is conducted on the authenticator side merely to trigger the secure element to do the specific cryptographic action. The biometric matching never leaves the authenticator. In this case, the whole authentication flow was divided into several isolated modules, which makes it easier to maintain and enhance the overall reliability.
- User to device authentication: Typically, if a user wants to interact with an IoT device that does not have a user interface, they need a primary interaction device or authenticator. The authenticator is required to register with the IoT device in the initial phase. The IoT device should have physical buttons that can be pressed and allow a specific time window for the authenticator to complete the registration. After a user verification on the authenticator, the scoped key pair—which is limited to only being recognizable by the specified IoT device—will be used to sign the challenge and output the response to the IoT device for signature verification. From the perspective of the IoT device, it only checks whether the authenticator is genuine or not. This method introduces two advantages: the first is convenience—devices can be plugged in using any verification method to verify the user. Also, it is easy to assure that the private keys are securely stored in secure elements and can be used only after verifying the user. The second is security—IoT devices can be compromised, so it is not suitable to use bearer credentials, e.g., a PIN or password, with the IoT device. Instead, by leveraging the public key cryptography, only the public key is exposed, and thus the method is not harmful to the user and the authenticator.
- Device to device authentication: In a similar fashion to User to Device authentication, an IoT device can establish a connection with other devices or cloud-connected devices and transfer messages from time to time. Unlike a primary interaction device that requires user verification to unlock key storage each time, a standalone device, e.g., an air conditioner, does not need user interaction. Therefore, a specialized silent authenticator that never demands any user interaction can be embedded in the device to perform essential operations securely.
- Device to cloud authentication: Some IoT devices cooperate with cloud services, for example, parcel lockers, hotel door entry systems, and rental cars. Consider a hotel online booking service that wishes to grant hotel room access to a specific user, or a logistics company wanting to allow a parcel recipient to open a particular locker. In general, a user will likely use online services for management and then go to a specific IoT device to perform further actions. Thus, we will have linked IoT devices and cloud services, with the cloud service's trust anchor recognizing the IoT devices. That is to say, IoT devices will connect to the cloud service and be capable of verifying the signature from the cloud service. This practice will require the service provider to embed the trust anchor into the device in the first place and let each user register with the cloud service using FIDO or other authentication standards. Later, the user will authenticate him-/herself via the cloud service and receive the response data with his/her public key associated with the cloud service's signature using proof of possession of the key in the authentication phase. This

proof of possession token will be used to authenticate the IoT device. This is different from bear tokens because a user will need to provide proof of possession of the private key in the course of authentication with the cloud service.

3. The Proposed Authentication System

The proposed authentication system includes three participants: the user, the authentication server, the IoT devices. The user is wearing a brainwave headset while operating IoT devices. The brainwave headset transmits user's brainwave data to the authentication server. The authentication server then authenticates the user by sending brainwave data into the authentication model and getting a result. If the current user is valid and authorized, the authentication server then gives IoT the permission to perform actions. Figure 2 presents the whole architecture.

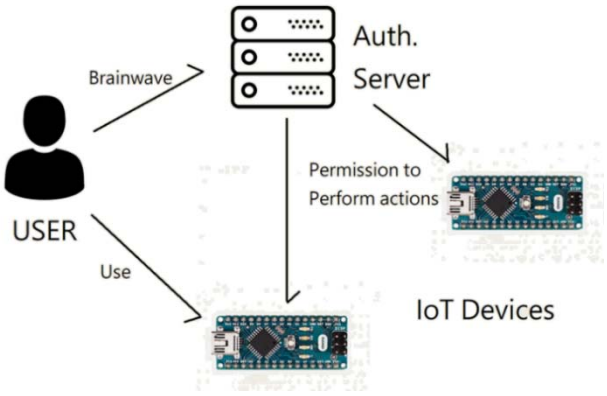

Figure 2. Overview of the proposed system.

In this research, we focus on the brainwaves corresponding to reactions between human memories and events that they have encountered, and we use a classifier to extract tokens of each individual from his/her brainwaves. In this way, we use the extracted tokens to create a brainwave-based authentication system. In the previous section, we mentioned that every individual has his/her own experiences of the environment. Although it is a dauntingly complex task to describe or quantify personal experience, there is still a way to do it. A person might have numerous divergent reactions or feelings toward different events. Although the reactions or feelings can vary, they can still be roughly categorized as "Familiar," "Déjà vu," and "Unfamiliar" [26]. If we collect images that include ones familiar to two individuals, and mix them with other unrelated images, we expect the two individuals' reactions would be as shown in Table 1.

Table 1. The recognition table of two individuals. A stands for images that P1 is familiar with, while B stands for images that P2 is familiar with. C stands for unrelated images for both.

Ind./Img.	A1	A2	B1	B2	C1	C2
P1	Familiar	Familiar	Unfamiliar	Familiar	Déjà vu	Unfamiliar
P2	Unfamiliar	Déjà vu	Familiar	Familiar	Unfamiliar	Déjà vu

If the above assumption is correct, we can expect that everyone has his/her own distinct familiarity pattern regarding a chain of images, based on his/her recognition pattern toward the environment. The more images we provide, the lower the probability of two individuals sharing the same familiarity pattern.

What can we expect from the familiarity patterns of individuals? Brainwaves are signals, external manifestations of an active brain and how it is functioning. Any action originating in the human brain has a corresponding brainwave from the very portion of the brain that is in charge of the specific action being signaled. For example, when you close your eyes, we can observe a slightly shivering

alpha wave embedded in the brainwave. Based on this result, we extended the theory to encompass familiarity. Specifically, we posit that an individual's familiarity with images might have a different corresponding brainwave. Based on the assumptions we have outlined, we designed our experiment as shown in Figure 3.

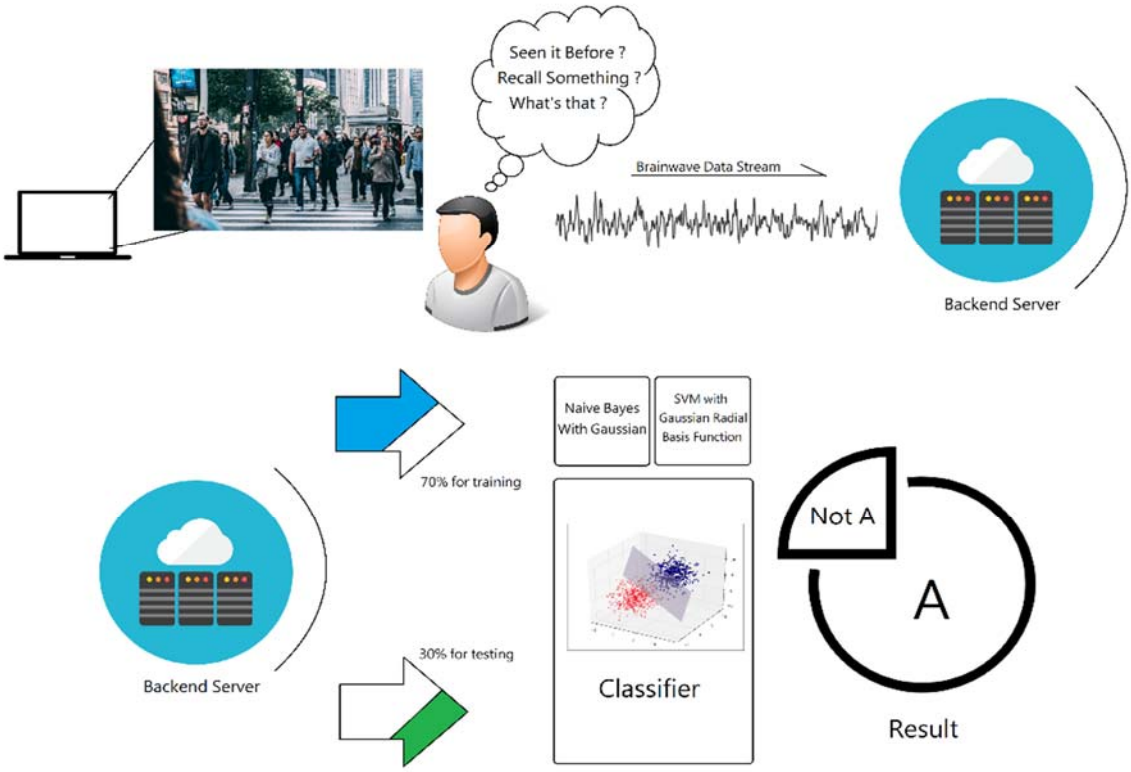

Figure 3. Overview of the experiment.

For the experiment, we prepared three different sets of images for the participants. For most of the image sets, these are images that particular participants are familiar with. Participants prepared these image sets. Every image was given an ID, and each image set was given a label associated with individual participants. We also stipulated to participants that the images they provided to us should not be exchanged with or otherwise sent to other participants. The last two sets of images are déjà vu and unfamiliar. Unlike the previous image sets, these are the only two image sets we provided. The déjà vu image set is images containing surroundings that participants are all familiar with. Since our participants are students on the same campus, we randomly took photos either on campus or in the immediate area around the campus. The unfamiliar image set contains images the participants are not familiar with. We collected these images by randomly searching for images of unpopular topics on the Internet (e.g., some street views of rural places). The reason for us avoiding pictures of popular topics (e.g., celebrities) is that these pictures have a higher chance of being seen by our participants, which may lead to some unwanted or unexpected effect of this image set.

With all the image sets collected, we scheduled times for the participants to take part in the experiment. The participants were asked to wear the brainwave headset (BR8) as they observed a series of images on the computer screen (see Figure 4). While displaying images on the screen, the computer sent the brainwave data it received to the backend server for further analysis. The program used to display the images does not require any interaction with participants, and all unnecessary elements on screen (i.e., icons, notifications) were turned off or eliminated to lower the number of possible extraneous influences (see Figure 5). Furthermore, to be sure no participants saw the images beforehand, participants who were waiting for their turn to experiment were asked to remain in a

different room from participants who were engaged in the experiment. Moreover, we made sure that participants who finished the experiment left without returning to the waiting room.

Figure 4. Actual view of the experimental setup.

Figure 5. Actual view of the experimental setting.

After the experiment, we retrieved data from the backend server for analysis. We took 70% of the data to train the classification model and left 30% of the data to test the trained data model. We use a BRI BR8 brainwave headset to collect brainwave data, an TravelMate 4750 Laptop to interact with participants, and an BM6AF PC as a server to store and perform data processing. The program that interacts with the participants is written in Java. BRI, the program bundled with the brainwave headset, collects brainwave data. Eltima Virtual Serial Port is the communication interface between the BRI program and the program interacting with participants. For data classifiers, we choose Support Vector Machine (SVM) as our classifier for the following reason. SVM works well on a wide range of classification problems, even problems in high dimensions that are not linearly separable [27]. We observed that brainwave data are both high-dimensional and streaming. Therefore, we think that SVM will be suitable for brainwaves.

BRI supports signal retrieval from time-based events. This means that, while the data are being collected, the other program (such as the program interacting with the participants) can send signals to BRI and BRI will make marks on the very record it is receiving at the same time. The signals can be varying; programs can send different kinds of signals based on events. They help a lot with data analysis. For example, a program can capture how participants react to a math problem. While a math problem was selected and given by the program to display on screen, the program sends a character "G" as signals to BRI; if the user clicked on the right answer, the program sends the character "R," or

otherwise sends the character "F." BRI stores brainwave data in CSV format. The first nine columns are timestamps and readings from Channels 1 to 8, and the last column is the event column. For our example, the raw data will look much more like Figure 6 displays:

1	21:22:29	0	0	0	0	0	0	0	0	71
2	21:22:30	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	0
3	21:22:31	-0.08	-0.08	-0.08	-0.08	-0.08	-0.08	-0.08	-0.08	0
4	21:22:32	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	0
5	21:22:33	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	0
6	21:22:33	-0.71	-0.71	-0.71	-0.71	-0.71	-0.71	-0.71	-0.71	0
7	21:22:34	-1.14	-1.14	-1.14	-1.14	-1.14	-1.14	-1.14	-1.14	82
8	21:22:35	-1.7	-1.7	-1.7	-1.7	-1.7	-1.7	-1.7	-1.7	71
9	21:22:36	-2.41	-2.41	-2.4	-2.41	-2.41	-2.41	-2.41	-2.4	0
10	21:22:37	-3.27	-3.27	-3.27	-3.27	-3.27	-3.27	-3.27	-3.27	0
11	21:22:38	-4.3	-4.3	-4.3	-4.3	-4.3	-4.3	-4.3	-4.3	0
12	21:22:39	-5.49	-5.49	-5.49	-5.49	-5.49	-5.49	-5.49	-5.49	0
13	21:22:40	-6.85	-6.86	-6.85	-6.86	-6.86	-6.86	-6.86	-6.85	0
14	21:22:40	-8.38	-8.38	-8.38	-8.38	-8.38	-8.38	-8.38	-8.38	70

Figure 6. BRI CSV file. (70, 71, and 82 are the ASCII coded of characters F, G, and R, respectively).

In this example, we can know when the computer gave the problem and when the participant reacted. Also, by using markings, we can identify which part of the data represents the time that the participant is thinking. We can see that the data between 71 and 82 represent a participant solving the problem and giving the right answer. It is beneficial that BRI provides such functionality; however, whether we can connect BRI with the program to receive event signals is another story. BRI supports RS232 (COM Port) as the only interface to receive signals. That is why we are using Eltima Virtual Serial Port as the communication interface.

We designed the participant interface to display images in the following pattern:

1. Calm down blank screen (15 s of blank screen for participants to calm down)
2. 5 unfamiliar images (each image displayed for 3 s, with a 3 s blank screen gap between images)
3. 5 déjà vu images (same display pattern as previous)
4. 5 familiar images (same display pattern as previous).

Between every change of the displayed image, the program sends a signal to the BRI program. The BRI program, at the same time, makes a mark every time the participant interface signals.

The 15 s of calm down time at the start is not only vital for ensuring participants have stable brainwaves but also necessary for the BR8 headset itself. Based on our observations, the headset also needs time to adapt. We provide brainwave graphs corresponding to two different time intervals for a person who is doing nothing but sitting still for comparison. In the graphs of the first 15 s (see Figure 7), we observe a huge spike, which is completely out of character compared with the graphs from 15 to 30 s (see Figure 8).

We carefully considered the time duration for displaying the images to ensure both enough time to measure brainwave reaction and enough time for the participant to remain clear-headed to prevent fatigue. A single round of the experiment took around 2 min, with 15 images displayed. The hardware capability of the BR8 headset allows it to sample brainwave signals at 1000 Hz. Therefore, 15 images with a duration of 3 s each can create 45,000 records of brainwave data.

BRI saves the brainwave data in CSV format, with event markers placed at the end of the specific data records. The participants' interaction program sends out character "G" while opening an image to display; while closing an image, it sends out character "C". These markers can assist us with pulling brainwave data that correspond only to the duration an image was displayed. We illustrate this in Figure 9.

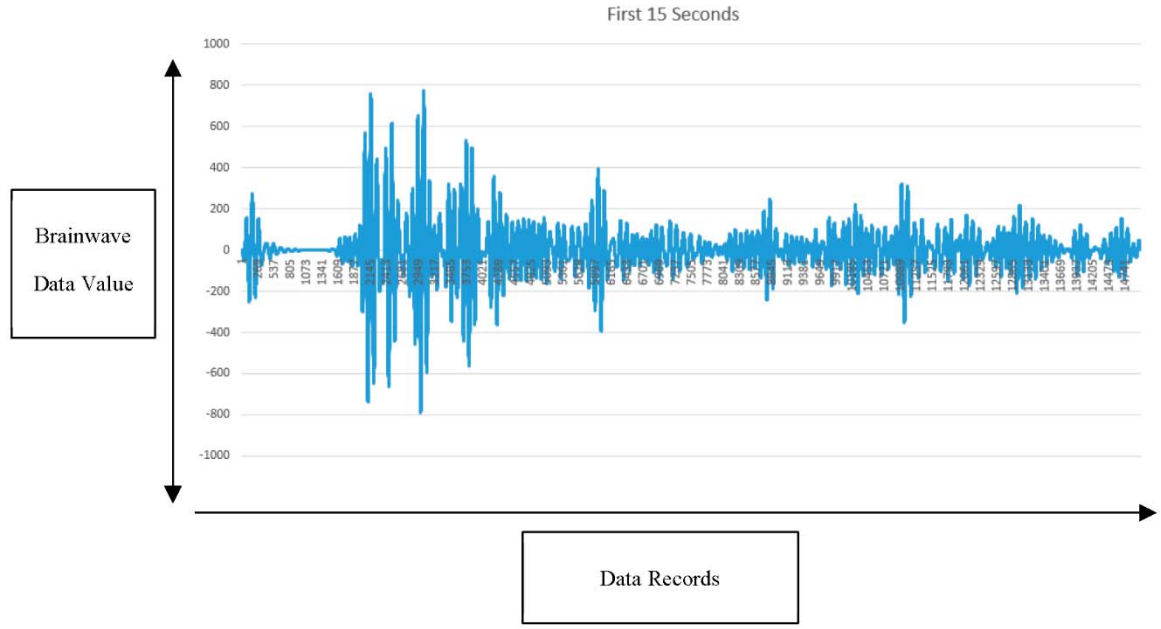

Figure 7. Graph of first 15 s.

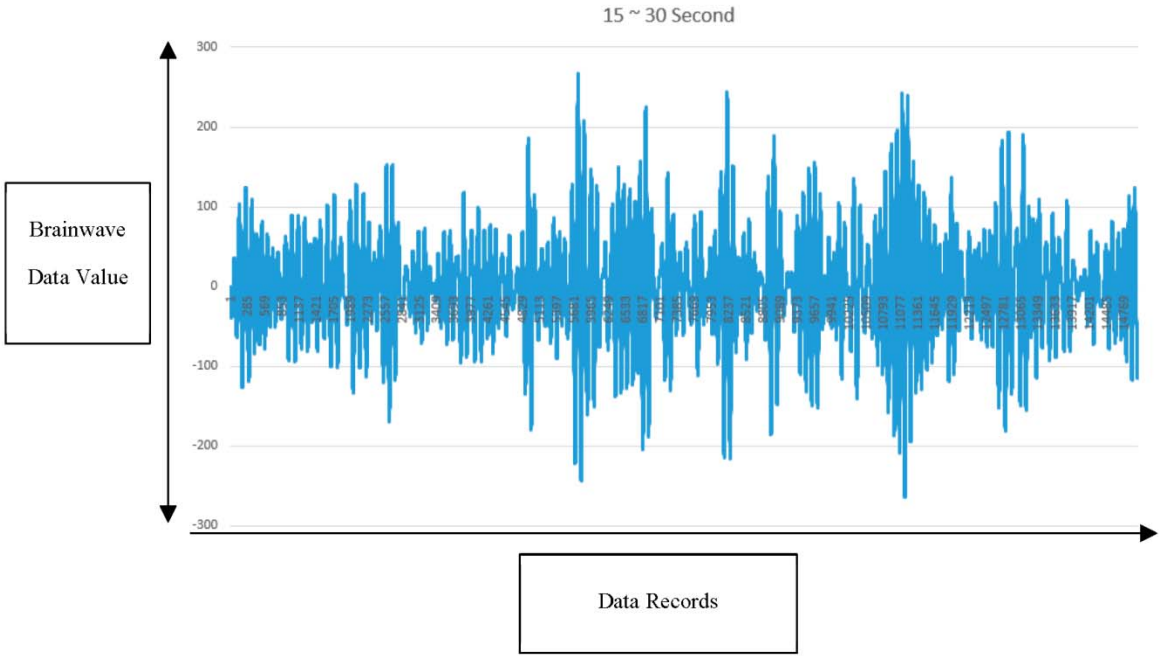

Figure 8. Graph of 15 to 30 s.

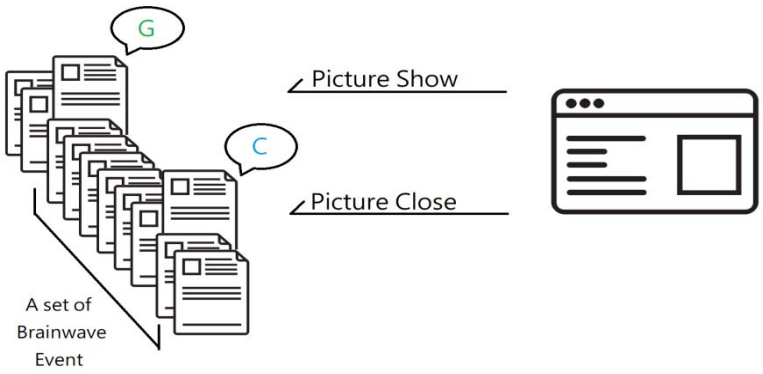

Figure 9. Illustration of how markers assist with creating a brainwave event dataset.

After the brainwave data collection process, we proceeded to the data preprocessing process. The brain has many different sections, each with different functionality. In this research, we focused on familiarity as an authentication token. To this end, the brainwaves originating from the parietal lobe are most likely the target data for us to analyze. If we look at the sensor placement of the BR8 headset, we can pinpoint brainwave data received by sensor Pz as the most interesting to us (see Figure 10) since this sensor is closest to the parietal lobe.

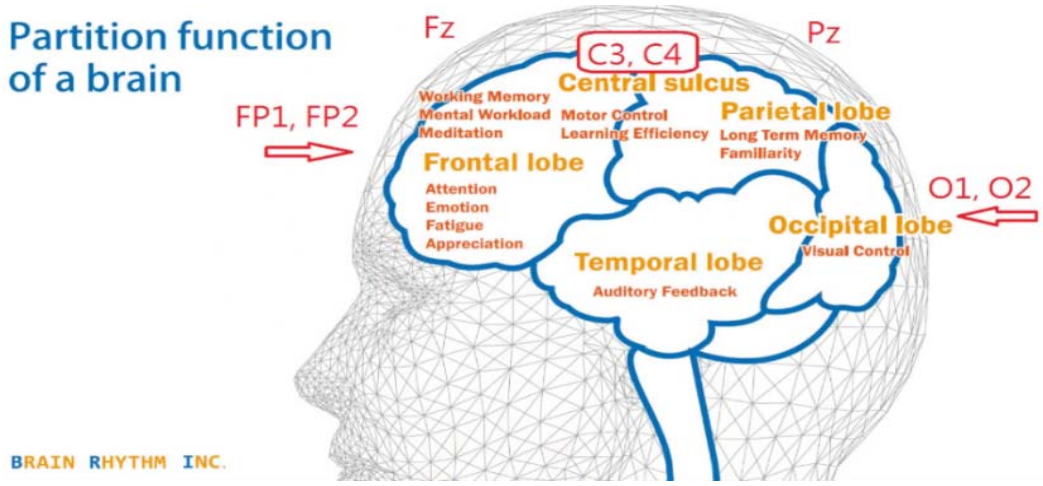

Figure 10. The sensor's locations and their corresponding sections of the brain [28].

In the early stage of data processing procedure tuning, we noticed an interesting phenomenon. The result of the classification by SVM is embarrassingly bad. We assume this is because the currently commercially available brainwave headsets are all non-invasive, relying on sensors being kept properly in contact to the participant's skin. Environmental issues can, therefore, influence the data the sensors obtain. For example, the skin condition of the participant (i.e., conductivity) or other electronic appliances in the immediate area (e.g., a high-power-consumption apparatus) can interfere with the headset's detection of brainwave signal data. Although the BRI software provides a notch filter for filtering out the 50 Hz/60 Hz noise caused by the alternating current running through the power cable, this is not enough. We applied Equation (1) to all the data to partially solve the issue and explain the behavior of the data. We use Figure 11 as an example for explanation.

$$\Delta R_i = R_i - R_{i-1} \tag{1}$$

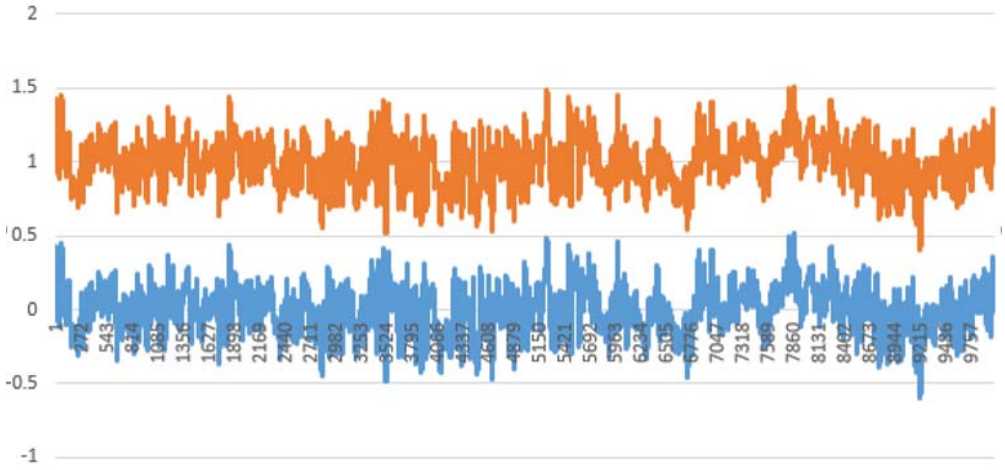

Figure 11. Example graph of brainwaves.

By applying the equation, which extracts only the delta value between data records, we can partially solve the interference of the environment. As we graph the handled datasets again, it appears as in Figure 12.

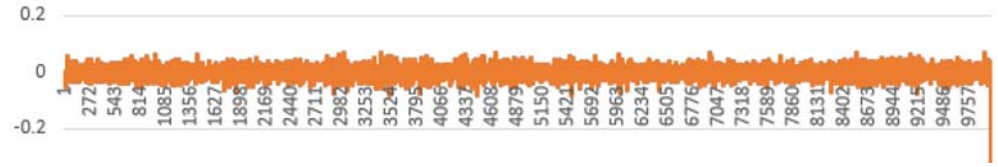

Figure 12. Example graph of brainwaves.

When we superimpose the two brainwave dataset graphs on each other, we can tell they are identical. When the equation is applied, the classification result of SVM improves significantly.

After completing the data preprocessing process, we started to filter out the similarities between profiles. We separated each participant's brainwave data into data for model training and data for model testing, at a ratio of 7:3. Then we appended all the data for model training from each participant to a single file and trained it into another model for similarity removal. We then likewise fed each participant's data in to be used for model training for similarity removal. This process ensured that we keep any correctly classified records in the data while discarding wrongly classified data. If any participant's data used for model training has 30% or more classified as others, the whole dataset of that particular participant would be eliminated and collected again. After removal of similarities, we trained the cleaned-up model training data into the model. We tested our model with the following procedure:

1. Choose a participant's data as a target.
2. Set the SVM label of other participants' data into the target participant to impersonate the target, to evaluate whether or not the classifier could correctly classify the testing data as others but not the target participant.
3. Feed the testing data into the trained model.

In brief, we expected the results illustrated in Table 2.

In this research, 30 participants took part in the experiment. That means there were 900 test cases. Similarity removal improved the results gradually. For comparison, we provide both the results of SVM classifier using the model data without similarities removed (Table 3) and the results of using the model data with similarities removal (Table 4) below.

Table 2. Brief test procedure illustration and expected result.

Test Set	Real Label	Faking as	Expected Result
1	A	A	Is A
2	B	A	Not A
3	C	A	Not A

In Table 3, we can tell that SVM wrongly classified 171 out of 900 cases, with FAR at 18.4% and FRR at 16.1%. In contrast, with similarities removed, we can see the slight improvement from Table 4: SVM only wrongly classified 130 out of 900 cases, with FAR at 13% and FRR at 13%. By running more rounds of similarity removal, the result will improve gradually. However, we do not recommend running many rounds of similarity removal. Applying such a method in too many rounds may lead to insufficient data. If the similarities between two profiles are significant, we suggest that the profile should be discarded or collected again.

We extracted the tokens from brainwaves successfully by using SVM after one round of data preprocessing and two rounds of similarity removal. If we took 30 participants' brainwave data into the system, it was capable of reaching 98.7% of accuracy with FAR at 1.1% and FRR at 1.1%, as shown in Table 5.

Table 3. Results of SVM classifier using model data without similarities removed.

Profile	Err. Count	Profile	Err. Count	Profile	Err. Count	Profile	Err. Count
CBH	1	HCH	8	LJJ	9	WJH	2
CYX	1	HJK	2	LRW	0	XCY	4
DCZ	3	HKU	2	LZZ	3	XKL	16
DOE	9	HSR	10	PYT	9	YRT	13
DOF	7	HSY	1	STU	10	YXU	1
DYH	4	HYL	3	SUT	2	ZYT	2
GJY	2	JYT	9	TAC	1		
GZY	15	LBY	2	TYX	20		

Table 4. Result SVM classifier using model data with similarities removed.

Profile	Err. Count	Profile	Err. Count	Profile	Err. Count	Profile	Err. Count
CBH	4	HCH	4	LJJ	12	WJH	6
CYX	1	HJK	1	LRW	1	XCY	0
DCZ	5	HKU	4	LZZ	4	XKL	2
DOE	11	HSR	9	PYT	9	YRT	5
DOF	13	HSY	4	STU	0	YXU	1
DYH	7	HYL	5	SUT	0	ZYT	0
GJY	2	JYT	11	TAC	1		
GZY	3	LBY	2	TYX	3		

Table 5. The final results with the authentication system.

Profile	Err. Count	Profile	Err. Count	Profile	Err. Count	Profile	Err. Count
CBH	0	HCH	0	LJJ	0	WJH	0
CYX	0	HJK	0	LRW	0	XCY	0
DCZ	1	HKU	0	LZZ	0	XKL	0
DOE	3	HSR	1	PYT	1	YRT	0
DOF	3	HSY	0	STU	0	YXU	0
DYH	1	HYL	0	SUT	0	ZYT	0
GJY	0	JYT	1	TAC	0		
GZY	0	LBY	0	TYX	0		

4. Conclusions & Future Work

In this research, we explore the link between experienced events and brainwave reactions in a specific area of the brain. By understanding the link between the encountered event and the area of the brain most likely to react to it, we can limit our focus to brainwave data from this specific part of the brain, eliminating potential unwanted data tuples in both the model and test data. Using a wearable headset with brainwave retrieval functionality built in and machine-learning classifiers, we successfully retrieved tokens from brainwave data and built an authentication system based on the tokens. The data pre-processing step affects the classification results of the SVM classifier dramatically. Without the pre-processing, the SVM classifier cannot correctly classify the data. After a similarity removal process, the results provided by the SVM classifier become more acceptable. The proposed system was capable of reaching 98.7% accuracy with FAR at 1.1% and FRR at 1.1%.

Although this paper pointed out the possibility of using brainwaves as an authentication token, there are more topics to discuss. In this paper, each participant only observed 15 images; this seems like a very small sample size for obtaining a reliable classifier. However, we still come up with acceptable results. It may seem to be as simple as adding more images, but that is not the only thing that needs to be considered. Fatigue and time are also issues to bear in mind. Adding more images means it takes longer for a participant to finish the experiment, which may cause fatigue—the participant may lose focus on viewing the picture. Instead, he or she would probably be thinking of something else, which affects the results. Also, another topic has been brought up: What if there are some changes in the participant's life experiences (such as becoming more or less familiar with something)? How can we make the classifier adaptive to marginal changes in the participant's life experience and automatically retrain the model? Finally, linking events and reactions in specific sections of the brain can provide us with a better view of the less relevant parameters hidden in brainwaves. Uncovering those links could allow us to exploit a wider range of possible uses of brainwaves.

Author Contributions: Conceptualization, W.C. and K.-H.Y.; Writing—original draft preparation, W.C., K.-H.Y., and C.-Y.F.; Writing—review and editing, W.C. and K.-H.Y.; Project administration, C.S., C.-M.C., and K.-H.Y.

Funding: This work was supported in part by JSPS Kakenhi Kiban(B) 18H03240 and Kakenhi Kiban(C) 18K11298, and in part by the Ministry of Science and Technology (Taiwan) under grants MOST 105-2221-E-259-014-MY3, MOST 105-2221-E-011-070-MY3, MOST 105-2923-E-182-001-MY3, and MOST 107-2218-E-011-012.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Musaddiq, A.; Zikria, Y.B.; Hahm, O.; Yu, H.; Bashir, A.K.; Kim, S.W. A Survey on Resource Management in IoT Operating Systems. *IEEE Access* **2018**, *6*, 8459–8482. [CrossRef]
2. Wurm, J.; Hoang, K.; Arias, O.; Sadeghi, A.-R.; Jin, Y. Security analysis on consumer and industrial IoT devices. In Proceedings of the 2016 21st Asia and South Pacific Design Automation Conference (ASP-DAC), Macau, China, 25–28 January 2016; pp. 519–524.
3. Zhang, Z.-K.; Cho, M.C.Y.; Wang, C.-W.; Hsu, C.-W.; Chen, C.-K.; Shieh, S. IoT Security: Ongoing Challenges and Research Opportunities. In Proceedings of the 2014 IEEE 7th International Conference on Service-Oriented Computing and Applications, Matsue, Japan, 17–19 November 2014; pp. 230–234.
4. Google SmartLock. Available online: https://get.google.com/smartlock/ (accessed on 12 April 2018).
5. How to Take Advantage of the Dynamic Lock Feature in Windows 10. Available online: https://www.techrepublic.com/article/how-to-take-advantage-of-.the-dynamic-lock-feature-in-windows-10/ (accessed on 12 April 2018).
6. Yeh, K.-H.; Su, C.; Chiu, W.; Zhou, L. I Walk, Therefore I Am: Continuous User Authentication with Plantar Biometrics. *IEEE Commun. Mag.* **2018**, *56*, 150–157. [CrossRef]
7. Matsuyama, Y.; Shozawa, M.; Yokote, R. Brain Signal's Low-Frequency Fits the Continuous Authentication. *Neurocomputing* **2015**, *164*, 137–143. [CrossRef]
8. Eye Closed Brainwave Dataset. Available online: http://www.bri.com.tw/data/sample_data/BR8_sample%20data_eyeclosed20141205.rar (accessed on 12 April 2018).

9. Abomhara, M.; Køien, G.M. Security and privacy in the Internet of Things: Current status and open issues. In Proceedings of the 2014 International Conference on Privacy and Security in Mobile Systems (PRISMS), Aalborg, Denmark, 11–14 May 2014.
10. Liu, Z.; Longa, P.; Pereira, G.; Reparaz, O.; Seo, H. FourQ on Embedded Devices with Strong Countermeasures against Side-Channel Attacks. In Proceedings of the Conference on Cryptographic Hardware and Embedded Systems 2017 (CHES), Taipei, Taiwan, 25–28 September 2017.
11. McGrew, D.A.; Viega, J. The Galois/Counter Mode of Operation (GCM). Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.694.695&rep=rep1&type=pdf (accessed on 20 July 2018).
12. Abed, F.; Forler, C.; Lucks, S. Overview of the Candidates in the CAESAR Competition for Authenticated Encryption. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.689.6253&rep=rep1&type=pdf (accessed on 20 July 2018).
13. Niruntasukrat, A.; Issariyapat, C.; Pongpaibool, P.; Meesublak, K.; Aiumsupucgul, P.; Panya, A. Authorization mechanism for MQTT-based Internet of Things. In Proceedings of the 2016 IEEE International Conference on Communications Workshops (ICC), Kuala Lumpur, Malaysia, 23–27 May 2016.
14. Hernández-Ramos, J.L.; Pawlowski, M.P.; Jara, A.J.; Skarmeta, A.F.; Ladid, L. Toward a Lightweight Authentication and Authorization Framework for Smart Objects. *IEEE J. Sel. Areas Commun.* **2015**, *33*, 690–702.
15. Ning, H.; Liu, H.; Yang, L.T. Aggregated-Proof Based Hierarchical Authentication Scheme for the Internet of Things. *IEEE Trans. Parallel Distrib. Syst.* **2015**, *26*, 657–667. [CrossRef]
16. Crawford, H.; Renaud, K. Understanding User Perceptions of Transparent Authentication on a Mobile Device. *J. Trust Manag.* **2014**, *1*, 7. [CrossRef]
17. Clarke, N. *Transparent User Authentication-Biometrics, RFID and Behavioural Profiling*; Springer: London, UK, 2011.
18. Lamkin, P. Wearable Tech Market to be Worth $34 Billion by 2020. 2016. Available online: https://www.forbes.com/sites/paullamkin/2016/02/17/wearable-tech-market-to-be-worth-34-billion-by-2020/#4d539bb33cb5 (accessed on 20 July 2018).
19. Tulceanu, V. Comprehensive brainwave authentication using emotional stimuli. In Proceedings of the 20th European Signal Processing Conference (EUSIPCO), Bucharest, Romania, 27–31 August 2012; pp. 1772–1776.
20. Liew, S.H.; Choo, Y.-H.; Low, Y.F. Fuzzy-Rough Nearest Neighbour classifier for person authentication using EEG signals. In Proceedings of the 2013 International Conference on Fuzzy Theory and Its Applications (iFUZZY), Taipei, Taiwan, 6–8 December 2013; pp. 316–321.
21. Lee, C.; Kang, J.-H.; Kim, S.-P. Methods of selecting electroencephalographic features for personal authentication. In Proceedings of the 2017 17th International Conference on Control, Automation and Systems (ICCAS), Jeju, Korea, 18–21 October 2017; pp. 1115–1119.
22. FIDO Alliance. Universal 2nd Factor (U2F) Overview. Available online: https://fidoalliance.org/specs/fido-u2f-v1.0-ps-20141009/fido-u2f-overview-ps-20141009.html (accessed on 20 July 2018).
23. Web Authentication: An API for Accessing Public Key Credentials Level 1. Available online: https://www.w3.org/TR/webauthn/ (accessed on 20 July 2018).
24. Bello, O.; Zeadally, S. Intelligent device-to-device communication in the Internet of Things. *IEEE Syst. J.* **2016**, *10*, 1172–1182. [CrossRef]
25. Hamilton, D. The Four Internet of Things Connectivity Models Explained. Available online: http://www.thewhir.com/web-hosting-news/the-four-internet-of-things-connectivity-models-explained (accessed on 20 July 2018).
26. Zhou, L.; Su, C.; Chiu, W.; Yeh, K.-H. You Think, Therefore You Are: Transparent Authentication System with Brainwave-oriented Bio-features for IoT Networks. *IEEE Trans. Emerg. Top. Comput.* **2017**, *99*, 1. [CrossRef]
27. SVM. Available online: http://www.nickgillian.com/wiki/pmwiki.php/GRT/SVM/ (accessed on 23 September 2018).
28. Brain Partition. Available online: http://www.bri.com.tw/ (accessed on 12 April 2018).