Seamless Payment System Using Face And Low-Energy Bluetooth

Yeongnam Chae[1] Kelvin Cheng[1] Pankaj Wasnik[2] Björn Stenger[1]
[1]Rakuten Institute of Technology, Tokyo
[2]Rakuten Institute of Technology, Bengaluru
{yeongnam.chae, kelvin.cheng, pankaj.wasnik, bjorn.stenger}@rakuten.com

Abstract

This paper introduces a multi-modal authentication approach for payments using a face image and the low energy Bluetooth (BLE) signal of a user's device. Devices of registered users transmit temporally changing one-time identifiers. During a payment request, the query face image is only matched with database entries corresponding to nearby users, thereby significantly reducing the complexity of the task. For cases in which a user does not carry their device, the system includes a fallback mechanism to PIN-based two-factor authentication. A classifier on depth data input is used to reduce vulnerability to presentation attacks. We conducted a user study of different payment methods and demonstrated our system at a public event with 951 users.

1. Introduction

Biometric identifiers are distinctive, measurable characteristics used to describe individuals [6], which can be categorized into physiological and behavioral properties [1]. We only consider physiological characteristics, which are related to the body, such as fingerprints, palm veins, iris patterns, or facial appearance. Of these, face recognition is a widely used approach, but it tends to have lower accuracy compared to other methods [8], see Figure 1.

The accuracy of face recognition models has improved continuously in recent years, and although the accuracy exceeds 99% on public datasets, it is still insufficient for authentication systems that require a high level of security. In the case of face-based payment systems, it is critical to avoid false acceptances. Another challenge is that face recognition from a large set of users is significantly more difficult than face verification, which is commonly used to unlock mobile devices. Many high-security systems therefore employ double authentication, prompting users to provide additional information, such as a PIN. While increasing the accuracy, two-factor authentication also increases the time required for the authentication process.

In this work, we propose a new approach for multi-modal authentication using the Bluetooth low energy (BLE) signal of mobile devices as the second authentication factor, which only requires the user to carry their device with them. For the case in which a user does not carry it, the system defaults to PIN-based authentication. We propose a non-collision PIN space, which is more secure than using a phone number. An anti-spoofing method using 3D face liveness detection guards against print and video attacks. We carried out a user experience study and report the performance during a large-scale event during which approximately one thousand users tested the system.

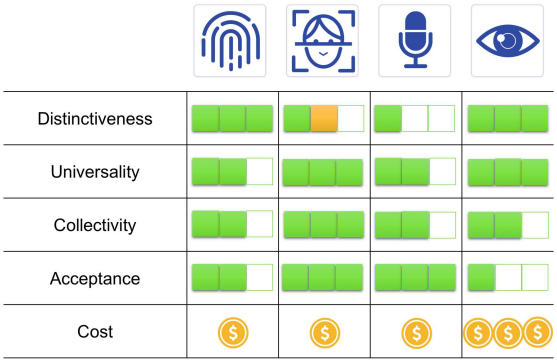

Figure 1. Comparison of different biometrics: fingerprint, face images, voice, and iris pattern.

2. Related work

Simple image-based face recognition systems are vulnerable to presentation attacks [3]. A large body of work exists on improving uni-modal systems by including secondary cues, such as the iris pattern [9], the user's voice [13], or the reflected pattern of an emitted sound signal [18]. The drawback of most of these methods is that an additional capturing device as well as user cooperation is required. Solutions using voice or sound rely on quiet environments to work robustly. Recently, devices with depth sensors are able to capture detailed 3D depth maps, achieving high accuracy for user verification [2]. One observed failure case includes

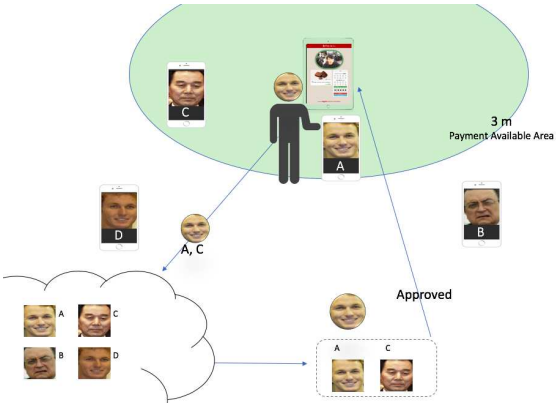

Figure 2. Payment system overview. During a payment request by user A, the payment kiosk extracts face features of A and BLE signal ids by users within the (green) payment area, A and C. The authentication server (bottom left) only needs to match against the database entries corresponding to these users.

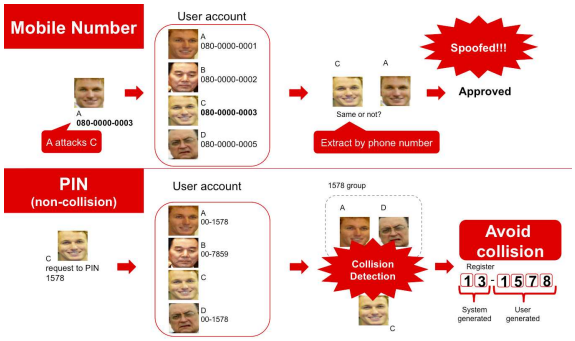

Figure 3. Comparison of secondary information. (Top) phone numbers are not always private, opening an attack vector, (Bottom) PINs selected by users may not be unique. Collisions can be avoided by adding a system-generated number.

that of identical twins, which may confuse the system.

3. System Overview

In a first registration step, a user's mobile device together with the user's face image is registered on an authentication server. The authentication system is integrated in a payment kiosk, which scans nearby mobile devices using the BLE signal. During a payment request, the BLE identifiers and face images are sent to the server via an Oauth2 secured channel. The authentication server extracts the locally encrypted face features from all sent identifiers and compares these with the query face features. If the query matches only one registered user, the authentication server approves the payment request. An overview of the proposed system is shown in Figure 2.

4. Bluetooth LE based authentication

In order to record device identifiers from nearby users, we use the advertising transmission signal of BLE and its received signal strength indication (RSSI). Bluetooth 4.0 LE includes advertising transmission for broadcast communication and operates at significantly reduced power levels compared to the original Bluetooth standard [7]. Since radio waves propagate according to the inverse-square law, distances can be approximated based on the relationship between transmitted and received signal strength (RSSI) as measured by a sensor [7].

4.1. Bluetooth LE Broadcasting

By transmitting temporal identifiers via the BLE advertisement channel, the authentication server can extract face information from its database without any additional user input. The security of the payment depends on the property of the secondary information. In case of a mobile number, it is unique, but not private. As shown in Fig. 3, a PIN is private but it may not be unique. However, the identifier sent by Bluetooth has both properties, being unique and private. The temporal identifier is refreshed periodically when users access the server, in order to prevent replay attacks. Furthermore, the advertising transmission can be used without pairing between central and peripheral devices and can be transmitted even while the application on the mobile device is in sleep mode. These characteristics make the proposed system work seamlessly from the user perspective.

4.2. Location-based candidate extraction via RSSI

By restricting the physical distance of a user from the payment kiosk via RSSI, the proposed system can significantly reduce the number of matching candidates from the face image as shown in Fig. 2. The distance between the kiosk and a user's device can be calculated by:

$$RSSI = -10n \log(d) + A \,, \tag{1}$$

where d is the distance, A the transmission power at $1m$ distance, n the signal propagation constant (normally $n = 2$). To decide the threshold of RSSI, we consider the distance and the signal propagation constant. The distance depends on the signal transmitting frequency and the averaging walking speed, estimated at $1.4m/s$ [15]. The signal propagation constant varies by room geometry [7]. We studied the Bluetooth signal strength under various conditions, *e.g.* when carried in a pocket or bag.

5. PIN-based secure authentication

Bluetooth based methods are seamless and easy to identify candidate users, but they need to carry their mobile device with them. Other face payment systems adopt mobile

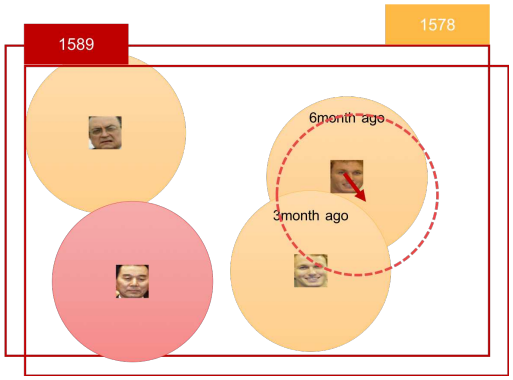

Figure 4. Lazy re-balancing of the PIN space: When a collision is detected, it prompts users to update their PIN.

numbers as secondary information [14], because it is unique and easy to memorize. However, it is not necessarily private and therefore vulnerable to phishing attacks. One alternative are PINs, which are private, but not necessarily unique. Collisions occur when users with similar appearance use the same PIN as shown in Fig. 3. In this paper, we propose two different approaches in order to avoid collision, by prevention and by re-balancing of the PIN space, respectively.

5.1. Collision prevention

The first approach to avoid collisions of user-selected PINs is to add a system-assigned PIN, such that users who may be confused by the system are assigned different PINs. As shown in Fig. 3, the proposed approach detects a collision when a user registers their PIN. In this case a system generated PIN (SGPIN) is determined by selecting the group of SGPIN with the largest minimum distance from the query face q:

$$s_q = {}_{s \in S} \left(\min_{x \in s} \delta(f_q, f_x) \right) \tag{2}$$

where s_q denotes the newly assigned SGPIN for query user q, S means set of SGPIN for the specific PIN that user requested, x denotes user included in a SGPIN group s, f represents facial feature which is extracted from the facial model, and δ denotes the distance of face feature vectors, typically Euclidean or cosine distance. This approach is secure by avoiding collisions, but it reduces convenience, because users need to memorize additional information.

5.2. Collision and lazy re-balancing

The second approach is to accept PIN collisions and use a third modality, such as a mobile number. To minimize this case occurring, we propose lazy re-balancing of the PIN space. When the system detects a collision, it prompts users who registered their PIN the longest time ago to refresh their PIN via a separate channel like e-mail, see Fig. 4.

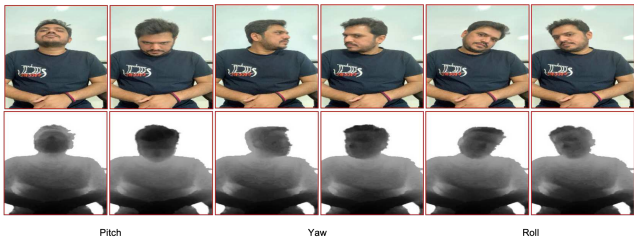

Figure 5. Sample face images for pitch, yaw and roll orientation.

By refreshing the collision PIN in a lazy manner, the entire PIN space is re-balanced over time.

6. Marginal threshold

In the case of collisions, which can occur during the collision of PINs or BLE signal, we request inputting the next-level information. We propose setting a marginal threshold for collision checking to reduce number of collisions as follows:

$$\delta(f_q, f_2) - \delta(f_q, f_1) < \theta_m \quad for\ \delta(f_q, f_1) < \delta(f_q, f_2) < \theta \tag{3}$$

where f_q denotes the query feature vector, f_1 and f_2 the closest and second closest candidate in the database, respectively, θ is the threshold decided by the face model, and θ_m denotes the marginal threshold. In the proposed system, we employ MTCNN [16] for the face detection and FaceNet [11] for the facial feature extraction. For the FaceNet, a recognition accuracy of 99.63% was reported on the LFW dataset [5]. Intuitively, a candidate is authenticated if there is a significantly closer match compared to other collision candidates.

7. Liveness detection

In order to defend against photo and video spoofing in a seamless manner, we use single shot 3D face liveness detection on a standard tablet with an embedded 3D depth sensor. 2D based approaches show limited performance than 3D based one [17]. We trained an Inception V3 model to determine the probability of an the input corresponding to a real face. Our training set contains 10-second clips of users at different distances to the camera, with and without head rotation, see Fig. 5. We collected 268 videos of 60 subjects in total. In order to collect negative samples, we simulated spoofing scenarios in which the attacker shows a photo or recording of a person to the camera at different distances, see Fig. 6. A total of 87 videos were collected this way.

To fine-tune the Inception V3 network, we removed the final fully connected layer and replaced it with a layer connecting to two classes, real and fake images. The network was trained using mini-batches of size 64 for 20 epochs using SGD with momentum (SGDM) [4]. We synchronize the

Figure 6. Depth map data helps guarding against presentation attacks. Positive (left) and negative sample (right).

RGB and depth data captured with different sensors. The RGB image is used to detect faces, and if a face is present, we pass the corresponding depth map region to the CNN to obtain the liveness score.

8. Experiments

8.1. BLE signal

In order to prepare the payment area, we measured the RSSI strength under different conditions: distance between the devices ($1m, 2m, 3m$), presence of obstacles (carried inside the pocket or outside the pocket), and the power mode of the BLE signal (low latency, balanced, low power). We used a tablet (Apple iPad Pro 11") as the payment kiosk and a phone (Samsung Galaxy S6 Edge) for the user's mobile device. For each combination, we collected 100 measurements. The power mode, A in Eq. 1, affects the overall RSSI level. We tested obstacle conditions, represented by n in Eq. 1, where the phone was inside or outside the pocket. We observed that being inside a pocket only minimally reduces the signal strength. A distinct signal strength difference inside vs. outside the pocket can be detected at short distances in low latency mode. However, we observed that signal strength was weaker at larger distances when other modes were used.

We also measured the transmitting interval of BLE advertisement across power modes and distances. By multiplying the average human walking speed with the average transmitting interval, we calculate the payment available area. Here we set the boundary of the payment area within a $3.2m$ radius from the terminal, which covers the Bluetooth low power mode. The RSSI threshold for the payment area was set to $-85dB$, which covers most of RSSI, including noise.

Power mode	*1m*	*2m*	*3m*
Low latency	0.10	0.11	0.13
Balanced	0.28	0.30	0.31
Low power	1.50	2.29	2.29

Table 1. Average transmitting interval of RSSI by distance and power mode (sec).

Type	*Precision*	*Recall*	*f1-score*
Real	0.94	0.99	0.96
Attack	0.98	0.88	0.92

Table 2. Performance metrics for 3D liveness detection.

8.2. Liveness Detection

The training of the Inception V3 network was carried out by learning the features of real and fake samples. Our training dataset consists of two classes with 70 frames selected from each real video and 120 frames from each spoof attack video. In total, the dataset consists of the $10,360$ real and $10,440$ fake samples, respectively. In the test phase, we used 10% of the frames of the whole dataset, resulting in 1,036 real and 1,044 fake comparisons. The face liveness detection results are presented using standard evaluation criteria *i.e.*, precision, recall and the f1-score. Table 2 presents details of the data and the obtained results. The model correctly identifies photo and video attacks and real samples, with recall of 0.88 and 0.99, respectively.

8.3. User study

We conducted a user study to evaluate our proposed approaches, as well as a previous approach where a phone number is used (similar to [14]). We asked users to register and try the various approaches in an imaginary scenario of purchasing chocolate. They were also asked to provide qualitative feedback in terms of ease of use, convenience, perceived security, and overall satisfaction.

The four conditions were:

(A) Payment using face and phone number,
(B) Payment using face and non-collision PIN (6-digit) approach,
(C) Payment using face and BLE signal approach,
(D) Payment using face and collision PIN (4-digit) approach and phone number.

A within-subject design was used so that all users completed all conditions once. The presentation order of the conditions was counterbalanced across all subjects by using a Latin Square. Each condition was conducted on a separate tablet device of the same model.

8.3.1 Participants

Thirty-four users (65% female) with an average age of 35 participated in the study. 12% had experience with face recognition payment system previously, 50% had experience with NFC payment, while 68% had previous experience with QR code payment. Around 21% have a technical occupations.

The study was reviewed and approved by our internal privacy department; written informed consent was agreed from each volunteer.

8.3.2 Procedures

Users first filled in a questionnaire with regard to demographic information, as well as their prior experience with new technology used for payment systems. They were then asked to register their face, phone number, and create a 4-digit PIN. This is done on a dedicated mobile device. They were then told a 2-digit system generated PIN, which was the same for all subjects. For condition B, this 2-digit PIN was combined with the 4-digit personal PIN to make it a 6-digit PIN, which was shown to the user on the registration device, giving them an opportunity to memorize it.

For each condition, users are shown a payment terminal in the form of a tablet device. They initiate the payment process by clicking on "Pay", and are then subsequently shown a camera screen where they can see their own image on the device, and face recognition will be performed. For condition A, the next screen prompts them to input their mobile number, after which payment is completed. For condition B, the next screen will ask them to input the 6-digit PIN. If they forgot their PIN, they are allowed two new input attempts, and if unsuccessful, their 6-digit PIN is revealed. For condition C, the user will be asked to hold the mobile device which they used for registering, simulating carrying their personal device. For condition D, the user will be asked for their 4-digit PIN. In certain cases, they may be asked for to input their phone number. After paying with each condition, users were required to rate the conditions using the 5-point Likert scale, stating their degree of agreement or disagreement with a set of questions. The questions were:

(1) The system is easy to use,
(2) The system is fast to use,
(3) The system is secure to use,
(4) Overall, I'm satisfied with using this as my daily payment option.

At the completion of the experiment, we also asked users to rank the conditions from most preferred to least preferred.

8.3.3 Equipment

The same Samsung Galaxy S6 Edge mobile device as in the previous experiment, was used to register users' faces, PIN numbers and mobile numbers. This mobile device was also used for condition C, where users had to imagine that this mobile device as belonging to them. A customized Android app was written specifically for this purpose. For each condition, a separate iPad Pro 11" tablet was used for testing by the users. Each tablet was loaded with a dedicated customized full-screen iPad app specific for each condition. Finally, another tablet was used for completing the pre and post experiment surveys.

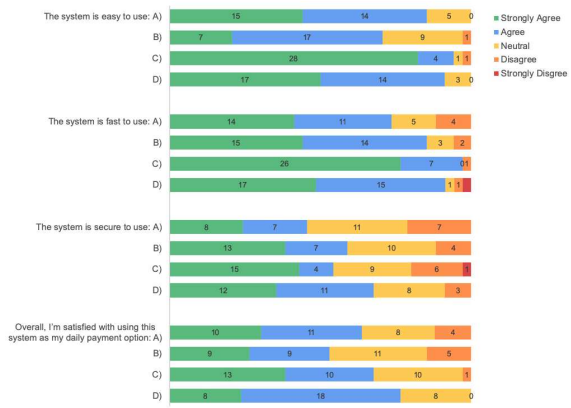

Figure 7. Users' agreement with each survey statement.

8.3.4 Results and Observations

Figure 7 shows a summary of the users' agreement with each statement. In order to analyze our survey data using parametric methods, we converted our ordinal data into interval values, as proposed by [12]. Hence, we used these equivalence: Strongly disagree = 1, Disagree = 2, Neutral = 3, Agree = 4, Strongly Agree = 5.

In terms of "easy to use", the median response for each conditions A, B, C, D were 4, 4, 5, 4.5 respectively. We can observe a majority of users strongly agreeing with this statement for condition C (28 users), compared to the other conditions. Indeed, a repeated measures ANOVA revealed a statistically significant difference between the various conditions F(3,135) = 10.177, p = 0.000. A post-hoc Tukey test reveals that condition C was significantly easier to use than conditions A and B at $p < 0.05$. Furthermore, condition B was significantly more difficult than all other conditions.

In terms of "fast to use", the median response for each conditions were 4, 4, 5, 4.5 respectively. We also observe that condition C has a higher number of users strongly agreeing with 26 users compared to other conditions. Again, a repeated measures ANOVA revealed a statistically significant difference between the various conditions $F(3, 135) = 5.432, p = 0.002$. A post-hoc Tukey test reveals that only condition C was significantly faster than conditions A and B at $p < 0.05$.

In terms of "secure to use", the median response for each conditions were 3, 4, 4, 4 respectively. A repeated measures ANOVA revealed a statistically significant difference between the various conditions $F(3, 135) = 2.875, p =$

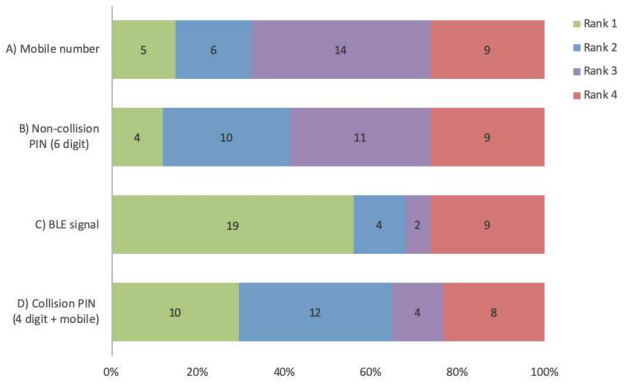

Figure 8. Ranking of the four conditions ordered by personal preference (#1 is most preferred, #4 least preferred).

0.040. A post-hoc Tukey test reveals that only condition A was significantly more secure than conditions D at $p < 0.05$.

In terms of overall satisfaction, the median response for each conditions were consistent, at 4 (agree). A repeated measures ANOVA revealed a statistically significant difference between the various conditions $F(3, 135) = 2.926, p = 0.038$. However, post-hoc Tukey test found no significance. This means that they were all quite satisfied with using all approaches equally.

We also gave users the opportunity to give us written feedback or comments for each conditions. In general, for condition A, the top concern for many users (7 users) was regarding security as their mobile number is known to others already. The other concern that 6 users had was that it took a bit longer to input 11 digits. Although some users thought it was actually quite convenient (5 users)

For condition B, the main concern was that they might not remember the 2 system generated PIN (10 users). On the other hand, 5 users thought it was actually quite easy to use. Additionally, we observed that many users took a photo of the 6-digit PIN that was shown on the registration device, with their own mobile device, in case they forget the PIN. Only in two cases where the users had actually forgotten their PIN, and didn't have them recorded anywhere, and we had to remind them what it was.

In condition C, users thought that it was fast or easy (7 users), yet a few users had concerns about the reliant of their phone, running out of battery, or having to remember to turn on Bluetooth. Surprisingly, two users thought that it happened way too fast, and that they might need some time to adjust to it. It was "so much easier than currently that it may take some getting used to. I feel like I might buy things accidentally or things I don't want. Basically so much quicker that I am taken aback."

For condition D, concerns were mostly about having to remember their PIN (4 users), and users questioned about the security of this method (4 users). One in particular had an issue about inputting their PIN on the tablet "I feel that to put my PIN on a big screen where everyone can see is a little scary."

As for face payment in general, one user in particular had concerns that the photo that they took was a bit "disappointing". In our registration app, the success screen shows the resultant photo which the system has captured. We found that several users cared how their photo looks in the system, even though it would not be shown to anyone else. This needs to be taking into account when designing face payment system to be deployed in public.

Figure 8 shows the overall results when users were asked to rank the 4 conditions from most preferred to least preferred. The median ranking for condition A to D are: 3, 3, 1, 2. We used the Friedman's Test and found a statistically significant difference between the ranking of the 4 conditions, $\chi^2(3) = 8.268$, p = 0.041. Post-hoc analysis with Wilcoxon signed-rank tests was conducted, and we found a significant difference between condition A and C (Z = -2.2603, p = .0238). Meaning that overall all participants, the general ranking is C, D, A&B (where A and B are third equal), with condition C clearly ranked higher than condition A.

8.3.5 Summary

In summary, from this user study, we found that the BLE signal approach has a higher general preference overall. It has the highest ranking, clearly out-ranking the phone number approach, as well as being preferred in terms of ease of use. It is also perceived to be faster than both phone number, and non-collision PIN approaches. However, due to it being a novel technique in terms of payment, general users have reservations, and may probably require some time to be familiar with such payment options, even if it is faster and easier to use.

Figure 9. Demonstration at a public event.

8.4. Demonstration at a large public event

In order to further verify the proposed system in more realistic situation, we demonstrated the proposed system in a large scale business conference. Overall, around 80,000 attendees attended the conference, from which 951 attendees tried the demonstration system during the four day event as shown in Fig. 9. In this demonstration, we adopted BLE signal and collision PIN and mobile number approaches in a hierarchical manner. During the demonstration, no false acceptances nor false rejections were recorded.

The demonstration and the study was reviewed and approved by our internal privacy department; written informed consent was agreed from each attendee.

8.4.1 Experiment on marginal threshold

In order to estimate the optimal marginal threshold, we analyzed the ratio of collisions and the approvals by secondary information. The overall result is shown in Table. 3. As shown in Table 3, we observed that 4.47% of all transactions of the BLE signal can be improved by the marginal threshold. In the case of PIN collision, around 20.42% of all transactions can be improved by the marginal threshold.

Type	*BLE Signal*	*4-digit PIN*
Collision	0.64	11.62
Approval by θ_m	4.47	20.42
Approval by θ	94.89	67.96

Table 3. Ratio of approvals and collisions by secondary information (%).

In Fig. 10, we plot the margin distribution for the transactions of multiple candidates. In this demonstration, we used the value of 0.2 as marginal threshold θ_m. As we can observe in Fig. 10, 0.2 covers around 80% of all transactions for multiple candidates. With the current marginal threshold, we did not observe any false acceptances nor false rejections, but further studies are needed to optimize the marginal threshold.

8.4.2 User experience feedback study

In order to further gauge user feedback on our proposed hierarchical approach, we invited conference attendees to try the proposed system and complete a short survey to further collect qualitative data. At this event, we have combined face with BLE signal approach, together with collision PIN or phone number. At this event, the procedure was similar to the user study that was described in the previous section. However, we did not have a pre-experiment survey. Similar to our previous study, we asked users to register their face, provide a 4-digit PIN, as well as their phone number. After

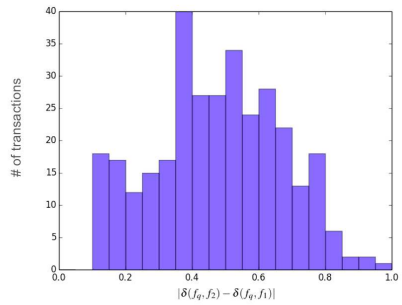

Figure 10. Margin distribution for the transactions of multiple candidates.

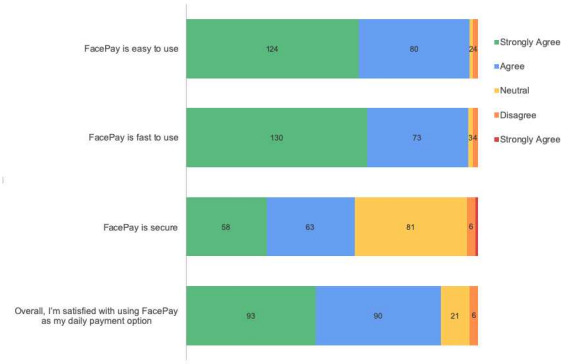

Figure 11. User experience feedback at a public event.

registration, we asked them to try our face payment system, condition C and D sequentially. We then asked users to provide feedback on ease of use, convenience, perceived security, and overall satisfaction, with similar questions as the previous user study. The equipment used were also the same.

210 completed surveys were collected from the 951 attendees who tried our system. Figure 1 shows the results from this survey. The median response for each question were: Strongly Agree, Strongly Agree, Agree, Agree, respectively.

As can be seen from these results, 97% of users found our system easy and fast to use. Perceived security was also generally agreed to by users (58%). Users are also positive about using such kind of system as their daily payment system with 87% of users agreeable or above.

9. Conclusions

This paper proposed a novel approach for a secure and seamless payment system by using the user's face image and the Bluetooth signal of their device. By broadcasting a temporal identifier via the advertisement transmission of BLE without pairing, we estimate a user's physical loca-

tion by the RSSI. We studied the relationship between signal strength, distance, obstacle and transmission interval, and based on this, we designed a payment-enabled area that identifies candidate users. We also proposed alternative PIN based methods as a fallback solution for users not carrying their registered device. The proposed system can improve overall accuracy and false acceptance ratio, which is critical for high security. By adopting a marginal threshold for collision detection we showed that we can further improve the accuracy. 3D face liveness detection is integrated and provides a necessary condition before authentication can proceed.

We conducted a user study with 34 volunteers to compare 4 different face payment approaches using a combination of other secondary inputs. The proposed face + BLE approach was shown to have a higher ranking reference, easier to use, and faster to use when compared to face + mobile number approach.

We also successfully demonstrated our system at a public event with 951 attendees with no false acceptances or false rejections. Current limitations include that the BLE signal transmission requires power, however, this is in the range of $0.1mW$ when transmitting the signal at $1s$ [10]. For future work, we will investigate secure BLE communication to prevent BLE signal replay attacks.

References

[1] *Handbook of Biometrics*. Springer, 2010.

[2] Apple. About face id advanced technology. `https://support.apple.com/en-us/HT208108`, 2018. [Online; accessed 18-September-2019].

[3] Nguyen Minh Duc and Bui Quang Minh. Your face is not your password face authentication bypassing lenovo–asus–toshiba. *Black Hat Briefings*, 4:158, 2009.

[4] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[5] Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, October 2007.

[6] Anil Jain, Lin Hong, and Sharath Pankanti. Biometric identification. *Commun. ACM*, 43(2):90–98, Feb. 2000.

[7] M. Kaczmarek, J. Ruminski, and A. Bujnowski. Accuracy analysis of the rssi ble sensortag signal for indoor localization purposes. In *2016 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pages 1413–1416, Sep. 2016.

[8] Salil Prabhakar, S. Pankanti, and Anil Jain. Biometric recognition: Security and privacy concerns. *Security Privacy, IEEE*, 1:33 – 42, 04 2003.

[9] Kiran B Raja, Ramachandra Raghavendra, Martin Stokkenes, and Christoph Busch. Multi-modal authentication system for smartphones using face, iris and periocular. In *2015 International Conference on Biometrics (ICB)*, pages 143–150. IEEE, 2015.

[10] Raphael Schrader, Thomas Ax, Christof Röhrig, and Claus Fuhner. Advertising power consumption of bluetooth low energy systems. *2016 3rd International Symposium on Wireless Systems within the Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS)*, pages 62–68, 2016.

[11] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 815–823, 2015.

[12] Gail M. Sullivan and Anthony R. Artino. Analyzing and interpreting data from likert-type scales. *Journal of Graduate Medical Education*, 5(4):541–542, 2013.

[13] Philip Tresadern, Chris McCool, Norman Poh, Pavel Matejka, Abdenour Hadid, Christophe Levy, Tim Cootes, and Sebastien Marcel. Mobile biometrics (mobio): Joint face and voice verification for a mobile platform. *IEEE pervasive computing*, 2012.

[14] The Verge. KFC in china tests letting people pay by smiling. `https://www.theverge.com/2017/9/4/16251304/kfc-china-alipay-ant-financial-smile-to-pay`, 2017. [Online; accessed 18-September-2019].

[15] Wikipedia. Walking — Wikipedia, the free encyclopedia. `http://en.wikipedia.org/w/index.php?title=Walking&oldid=913600845`, 2019. [Online; accessed 14-September-2019].

[16] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10):1499–1503, Oct 2016.

[17] Shifeng Zhang, Xiaobo Wang, Ajian Liu, Chenxu Zhao, Jun Wan, Sergio Escalera, Hailin Shi, Zezheng Wang, and Stan Z. Li. A dataset and benchmark for large-scale multi-modal face anti-spoofing. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[18] Bing Zhou, Jay Lohokare, Ruipeng Gao, and Fan Ye. Echoprint: Two-factor authentication using acoustics and vision on smartphones. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, pages 321–336. ACM, 2018.