ICWMC 2023

The Ninth International Conference on Wireless and Mobile Communications

ISBN: 978-1-68558-060-5

March 13th - 17th, 2023

Barcelona, Spain

ICWMC 2023 Editors

Sonia Ben Rejeb, Higher Institute of Computer Science (ISI), University of Tunis El Manar (UTM) Tunisia

ICWMC 2023

Forward

The Nineteenth International Conference on Wireless and Mobile Communications (ICWMC 2023), held between March 13th and March 17th, 2023, continued a series of events on advanced wireless technologies, wireless networking, and wireless applications.

ICWMC 2023 addressed wireless related topics concerning integration of latest technological advances to realize mobile and ubiquitous service environments for advanced applications and services in wireless networks. Mobility and wireless, special services and lessons learnt from particular deployment complemented the traditional wireless topics.

We take here the opportunity to warmly thank all the members of the ICWMC 2023 technical program committee, as well as all the reviewers. The creation of such a high-quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and effort to contribute to ICWMC 2023. We truly believe that, thanks to all these efforts, the final conference program consisted of top-quality contributions. We also thank the members of the ICWMC 2023 organizing committee for their help in handling the logistics of this event.

We hope that ICWMC 2023 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the field of wireless and mobile communications.

ICWMC 2023 Chairs

ICWMC 2023 Steering Committee

Dragana Krstic, University of Niš, Serbia
Rajat Kumar Kochhar, Ericsson, Sweden
Magnus Jonsson, Halmstad University, Sweden
Sonia Ben Rejeb, Higher Institute of Computer Science (ISI), University of Tunis El Manar (UTM), Tunisia

ICWMC 2023 Publicity Chairs

José Miguel Jiménez, Universitat Politecnica de Valencia, Spain
Sandra Viciano Tudela, Universitat Politecnica de Valencia, Spain

ICWMC 2023
Committee

ICWMC 2023 Steering Committee

Dragana Krstic, University of Niš, Serbia
Rajat Kumar Kochhar, Ericsson, Sweden
Magnus Jonsson, Halmstad University, Sweden
Sonia Ben Rejeb, Higher Institute of Computer Science (ISI), University of Tunis El Manar (UTM), Tunisia

ICWMC 2023 Publicity Chairs

José Miguel Jiménez, Universitat Politecnica de Valencia, Spain
Sandra Viciano Tudela, Universitat Politecnica de Valencia, Spain

ICWMC 2023 Technical Program Committee

Mohamed Abid, University of Gabes, Tunisia
Bedoui Abla, CEDOC-2IT | INPT (National Institute of Posts and Telecommunication), Morocco
Afrand Agah, West Chester University of Pennsylvania, USA
Iness Ahriz, CNAM, France
Khalil Aissaoui, Tunisia Polytechnic School (TPS), Tunisia
Wafa Akkari, University of Manouba, Tunisia
Ali Kadhum M. Al-Quraby, University of Babylon, Iraq
Diego Alberto Godoy, Universidad Gastón Dachary, Argentina
Adel Aldalbahi, King Faisal University, Saudi Arabia
Adda Ali-Pacha, University of Sciences and Technology of Oran, Algeria
Karine Amis, IMT Atlantique, France
Tran Hai Anh, Hanoi University of Science and Technology (HUST), Vietnam
Antonio Arena, University of Pisa, Italy
Kamran Arshad, Ajman University, UAE
Radu Arsinte, Technical University of Cluj-Napoca, Romania
Nebojša Bačanin-Džakula, Singidunum University, Serbia
Salih Safa Bacanli, University of Central Florida, USA
Nedia Badri, ENSI - University of Manouba, Tunisia
Corey E. Baker, University of Kentucky, USA
Chaity Banerjee, University of Central Florida, USA
Kamel Barkaoui, Cedric | Cnam, France
Paolo Barsocchi, ISTI (Institute of Information Science and Technologies) | Italian National Research Council (C.N.R.), Pisa, Italy
Hadda Ben Elhadj, SM@RTS | Higher Institute of Informatics | Monastir University, Tunisia
Sonia Ben Rejeb, Higher Institute of Computer Science (ISI) - Higher School of Communications of Tunis (SUPCOM), Tunisia
Emna Ben Slimane, ENIT - Tunis El Manar University, Tunisia
Djamila Bendouda, Ecole Nationale Supérieur de Technologie, Algeria
Driss Benhaddou, University of Houston, USA
Djedjiga Benzid, École de Technologie Supérieure - Université du Québec, Canada

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

A Sound Events Detection and Localization System based on YAMNet Model and BLE Beacons

Carlos M. Mesa-Cantillo[†], Itziar Alonso-González[†‡], Miguel A. Quintana-Suárez[‡], Carlos Ley-Bosch[†‡], Carlos Ramírez-Casañas[†‡], Javier J. Sánchez-Medina[*], David Sánchez-Rodríguez[†‡]

[†]Institute for Technological Development and Innovation in Communications, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain, email: {carlos.mesa, itziar.alonso, carlos.ley, carlos.ramirez, david.sanchez}@ulpgc.es

[‡]Telematic Engineering Department, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain, email: mangel.quintana@ulpgc.es

[*]CICEI, Innovation Center for the Information Society, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain, email: javier.sanchez@ulpgc.es

***Abstract*—Automatic sound event detection is a must-have feature for several emerging applications, such as surveillance, automatic listening, and noise source identification. Acoustic Event Detection (AED) aims to know the sounds' identity and temporal position in signals captured by one or several microphones. In this work, we use a pre-trained Yet Another Mobile Network (YAMNet) model to perform real-time audio classification. That audio event classifier model takes the audio waveform as an input and makes independent predictions for each of the 521 audio events in the AudioSet ontology. The model used the MobileNet v1 architecture and was trained using the AudioSet corpus. By means of a Raspberry Pi 3, a commercial microphone, and a set of Bluetooth Low Energy (BLE) beacons, this system is able to detect potentially harmful events. Thus, the system can detect where and what event has been detected and send it to a database. After that, the database is updated, and a notification can be sent to the users of a specific application. This information may be helpful for people with disabilities so they can be warned of danger in the nearby areas.**

***Index Terms*—sound event classification, beacons, machine learning, yamnet.**

I. Introduction

The task of detecting the onsets and offsets of target class activities in general audio signals is known as Sound Event Detection (SED) [1]. It is a technique that accepts an audio signal as input and produces temporal activity for target classes like "vehicle passing by," "footsteps," "people chatting," "gunshot," etc. The time resolution of class activities might vary between techniques and datasets.

The use of audio sensors in surveillance and monitoring applications is especially useful for event detection. Detection systems can effectively notify when an event occurs while enabling further processing, such as notifying the system with information about the event and its position. In recent years, Acoustic Event Detection (AED) [2], [3] and Acoustic Event Classification (AEC) have been essential for many applications such as security surveillance [4], human-computer interaction [5], and "machine hearing" [6].

In most surveillance systems, locating the position of the acoustic source over a topological grid is the final objective of sound localization. In environments with little reverberation time, like a typical public square, the Time Difference of Arrivals (TDOA) of the signal at an array of microphones is the most used technique for source localization. These time delays are further analyzed in [7] to determine the source location.

On the other hand, indoor positioning technology is being commercialized in different technologies and qualities. Unlike the Global Positioning System (GPS), which is used for outdoor positioning, there is no proven method that can be used for all purposes in indoor positioning, at the moment of writing this paper. The available technologies employed nowadays may vary in terms of cost, accuracy, and maintenance requirements. Some of the most common may be the following:

- Bluetooth Low Energy (BLE): Signals from battery-powered beacons are the core of indoor location technology. It is one of the most remarkable technologies that have emerged for indoor location. It uses BLE beacons or iBeacons, which are cheap, small, have long battery life, and do not require an external power source. A receptor device can detect the beacon signal and can roughly calculate the distance to the beacon.
- Wi-Fi: it can be used similarly to BLE beacons but requires an external power source and, higher setup and hardware costs. The signal tends to be stronger than BLE, with a larger range.
- Ultra-Wideband: it is the most precise method for indoor positioning available. Nevertheless, compared to its alternatives, it has more hardware requirements, as well as higher costs.

In this paper, a system for sound event detection and localization is proposed. It is based on the Yet Another Mobile Network (YAMNet) model for sound detection and BLE beacons to infer the location. YAMNet is based on the Visual Geometry Group (VGG) architecture and employs the Mobilenet v1 depthwise-separable convolution architecture. The classifier has 3.7M weights and performs 69.2M multiplies for each 960 ms input frame. Although the YAMNet model has a total of 521 classes, only the ones that may be relevant for notification will be used. In [8], the accuracy of

 ISBN: ISBNFILL

this model is tested with only 6 out of the 527 classes of the Audioset database due to the simplification of the experiment, as well as time and computational constraints. YAMNet can classify single fixed-size audio samples with 92.7% accuracy and 68.75% precision.

In addition to the use of microphones for the SED, in this system, BLE beacons are used to associate each microphone with a specific location. This location is previously set in the database to correspond to the Media Access Control (MAC) of the beacon, so when the Raspberry Pi code scans for beacons, the beacons are filtered and only those beacons that have been configured in the database with a location associated with them will appear. This information associated with an application that can notify and warn about the environment can be helpful for people with disabilities. For example, in the case of someone who is deaf, a notification will appear on their phone screen and warn them that an event has occurred nearby and that they should be alert. This could help them avoid any possible danger. Also, since cell phones can read notifications, if the person is blind, an audio notification about a dangerous event that has occurred in a nearby area may be helpful for them.

This article is organized as follows. Section II summarizes the related work about sound event detection and indoor positioning systems. Section III briefly describes the hardware and software employed. Section IV introduces the methodology to start the real-time audio classification. Next, in Section V, we analyse the results of the working system. Finally, in Section VI, the conclusion and future work are presented.

II. Related Work

Audio tagging is the task of detecting the presence or absence of a certain sound event in a recording. This has several applications, such as surveillance, monitoring, and health care [9].

In many papers such as [10], [11], the input is a polyphonic sound in an undefined context, and the output is one identified sound event at each instance of the polyphonic sound. The sound event is determined by the most prominent one in that instance.

Polyphonic events are the main error source of AED. This problem is usually solved by treating the AED task as a multi-label classification problem. In [12], to better handle polyphonic mixtures, the authors propose to consider each possible label combination as one class. In other works, to handle event overlaps, the AED task is usually framed as a multi-label classification problem. This is solved by a deep neuronal network with multi-label output [13], [14].

In [15], a solution for the polyphonic SED task on mobile devices is presented. Its architecture includes offline training and online detection. The offline training involves model training and compression, and the online detection process consists of acquiring sensor data, audio processing, and detecting sound events.

Indoor positioning systems already have broad applications for providing localized information and directions. In [16], an indoor positioning system with room-level accuracy is proposed. It focuses on an installation procedure that non-technical staff can easily follow. In addition, it has a low cost.

Most of the existing solutions, employing beacons and smartphones, require the floorplan of the indoor environment and many beacons. These applications usually try to increase the accuracy of the coordinate estimation on Line Of Sight (LOS) environments using a large number of beacons [17] and large training datasets.

The most common methods used in indoor positioning systems are:

- Trilateration: the distance from the source to the receiver is used to estimate the user's location. A more detailed description of this method is presented in [18].
- Triangulation: a method for calculating a position that relies on a known distance between two measuring apparatuses and the measured angles from those two points to an object [19].
- Fingerprinting: this method consists of two stages. First, Received Signal Strength (RSS) measurements are captured for multiple points in the indoor environment. Then, these measurements are used to determine the user's location.

III. Materials

The hardware and software used to operate this system are as follows:

A. Raspberry Pi 3

A computer is required to run the classification script and send the information to the database. Raspberry Pi was selected because it is an inexpensive and compact computer that meets the software requirements for classification. This computer runs Raspberry Pi OS as its operating system and requires an Internet connection to communicate with the database. Also, Python 3.9 needs to be installed to execute the scripts with the required libraries.

B. iBKS 105

iBKS 105 is the BLE beacon that was employed. Its specifications are as described in Table I. These beacons are configured to work with the iBeacon protocol, with a transmission (Tx) Power of 0 dBm and an advertising interval of 1 second.

C. Krom Kimu Pro

The microphone used to capture the audio in real time was a commercial microphone, in this case, a Krom Kimu Pro. The specifications are as shown in Table II. This microphone was selected since it is a commercial microphone whose sample rate and the number of channels can be configured in the PyAudio library to correspond to the audio input of the YAMNet model.

D. Software

The code for the real-time classification was done in Python 3.9. The most relevant libraries that were employed are the following ones:

 ISBN: ISBNFILL

TABLE I
iBKS 105 SPECIFICATIONS

Size	11,3 x Ø52,6 mm
Battery lifetime	30-40 months (Tx power at 1s interval)
Protocols	iBeacon and Eddystone

TABLE II
KROM KIMU PRO SPECIFICATIONS

Sensitivity	$-35dB \pm 3dB$
Impedance	2.2K ohms
Frequency Response	20 Hz - 20 KHz
Sample Rate	48 KHz at 16 bits

1) PyAudio: To collect the real-time audio, the library PyAudio was selected since it allows configuring the sampling rate, the number of channels, the number of frames per buffer, and which microphone will capture the audio. The audio input must be configured as 16 kHz mono audio. This setting is required because it is the one supported by the YAMNet model.

2) TensorFlow Lite: TensorFlow Lite library was employed since the YAMNet model is a pre-trained model given by them. This model was trained with audio features computed as follows:

- All audio is resampled to 16 kHz mono.
- A spectrogram is computed using the Short-Time Fourier Transform magnitudes with a window size of 25 ms, a window hop of 10 ms, and a periodic Hann window.
- A mel spectrogram is computed by mapping the spectrogram to 64 mel bins covering the range 125-7500 Hz.
- A stabilized log mel spectrogram is computed by applying log(mel-spectrum + 0.001), where the offset is used to avoid taking a logarithm of zero.
- These features are then framed into 50%-overlapping examples of 0.96 seconds, where each example covers 64 mel bands and 96 frames of 10 ms each.

Additionally, since the computer that will support the real-time classification is a Raspberry Pi 3, and it cannot support some of the libraries that can be used on a desktop computer, it was necessary to use other libraries which require less processing power. In this case, the library employed was TensorFlow lite. The TensorFlow lite version of the YAMNet model has some additional changes:

- The model is quantized. The model is retrained with Relu6 non-linearities instead of Relu to limit the activation ranges.
- The inference signature is simpler. It takes a fixed-length audio frame and returns a single vector of scores for 521 audio event classes.

3) Bluepy: This is a project to provide an API to allow access to Bluetooth Low Energy devices from Python. At the moment, it runs only on Linux. This library has been used to scan the BLE beacons near our Raspberry Pi to speed up the assignment of the corresponding zone.

IV. METHODOLOGY

The methodology followed in this research work is presented in Fig. 1, which development is described as follows:

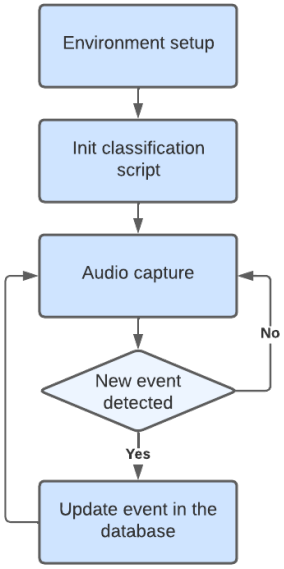


Fig. 1. The proposed methodology.

First, a previous environment setup is required. The beacons must be registered in the database and associated with a known location. This is made by designating each MAC of the beacons that will be employed to a known location of the indoor environment. Also, the iBKS 105 beacons need to be configured with their own application. In this application, the BLE service that will be used can be chosen between, Eddystone or iBeacon. The beacons were configured as iBeacon with a 1000 ms advertising interval and radio Tx Power of 0 dBm.

To register the beacons to the database, for 10 seconds a script developed in python scans for the nearest beacons. Then, it displays a list of the MAC addresses that were discovered with the name and the Received Signal Strength Indicator (RSSI) of the device as shown in Fig. 2. After that, the user can choose the desired MAC and select the region and the section where the beacon is displayed. The database structure will look as in Fig. 3. The audio is associated with an area or section instead of the beacon because multiple beacons or microphones can be associated with the same area. After the beacons are in their determined location, the Raspberry Pi and microphone pair can be placed near the desired beacon.

```
[?] Select your beacon: E2:FB:42:63:0A:67 - paloma5 - RSSI: -82
 > E2:FB:42:63:0A:67 - paloma5 - RSSI: -82
   6F:82:D7:DD:7B:EF - [LG] webOS TV LK610BPLB - RSSI: -78
   D4:A3:4A:E6:34:84 - Mi Smart Band 4 - RSSI: -97
   EE:E0:03:17:0A:54 - Mi Smart Band 6 - RSSI: -70
   other
```

Fig. 2. Adding beacon to the database.

 ISBN: ISBNFILL

Fig. 3. Database structure.

Once the Raspberry has been initialized, and the script runs, it will scan for nearby beacons. A list of the detected beacons that are registered in the database will be displayed, showing their MAC address, region, section, and RSSI. The user can choose from all the possible options for sound event detection to update the information in the database associated with that location. Next, the user can select the microphone to capture the audio in real-time. Once this configuration has been made, the script will start capturing audio. Whenever it detects a new event happening in relation to the previous one, it will update the database with the new event. The audio captured with the microphone must be 16 kHz mono audio to analyze sound events. This audio configuration is necessary for the YAMNet model to perform a correct audio classification. This adjustment was made when the microphone was selected.

V. Results

As mentioned in the previous section, the used components were configured to test the system and verify that it works correctly. First, several iBKS beacons were configured with the iBKS config tool application. Then, they were placed in different locations and registered in the database with the python script, so there were multiple options to set the microphone and Raspberry Pi.

Firebase was implemented as the back-end of the application, and the Firebase real-time database was used. This database is structured like a JavaScript Object Notation (JSON) file. In this database, the beacons were registered, as shown in Fig. 3.

Once the beacons were registered in the database, the script was run on the computer. Then one of the locations associated with the previously set beacons and the Krom Kimu Pro microphone was selected, after which the audio was captured.

YAMNet was re-trained in a transfer learning approach, for 5 out of the 50 categories in the ESC-50 dataset [20], to later test the accuracy of the model. Each of these categories consists of 40 sounds. The selected categories were the sounds of dogs, knocking on wooden doors, coughing, sirens, and vacuum cleaners. A total of 200 sounds were selected and divided into five different folds. Defining a very simple sequential model with one hidden layer and five outputs to recognize the sounds described before. For the new model, 60% of the samples were employed for testing, 20% were employed for validation, and the remaining samples were employed for testing. The model was configured to train for 20 epochs in the training phase. The accuracy obtained was around 87% in the 20 rounds that were done. In addition, a confusion matrix with 20 different sound samples was generated to visualize the results. The results obtained for these 20 samples are shown in Fig. 4.

To test the real-time update of the database, some sounds were emulated, such as knocking on a door, the sound of a siren, or dog sounds. These sounds were classified as intended, and the database was updated after a few seconds each time the sound changed. This update only rewrites the event of the previously selected location; if the event being played is the same, the database will not be updated.

VI. Conclusion and Future Work

In this work, it is described a system that is easy to implement and can classify real-time audio events and update a database with the detected location and associated information of such events. This system could be implemented in different environments to help people with disabilities. This assistance could take the form of an audible notification whenever a potentially dangerous event occurs near the user, requiring his or her attention.

In our ongoing work, we are planning to implement the system in a mobile phone application that detects where the user is located and warns them there might be a danger for a user with a disability. Also, the application could send push notifications so the user could be aware of the location

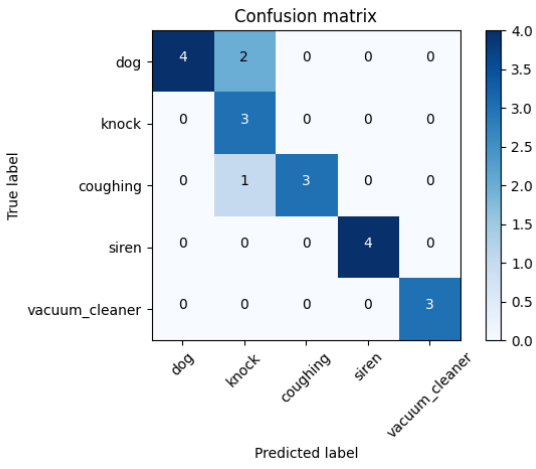


Fig. 4. Confusion matrix of the five classes for 20 different samples.

 ISBN: ISBNFILL

and the detected event. Additionally, with the implemented beacons, an indoor positioning system could be implemented in the application, enabling more accurate indications about the events that are being captured.

ACKNOWLEDGMENT

This research has been partially founded by the Consejería de Economía, Conocimiento y Empleo del Gobierno de Canarias, Agencia Canaria de Investigación, Innovación y Sociedad de la Información under projects ProID2020010009 and CEI2020-08, Spain.

REFERENCES

[1] E. Çakır, G. Parascandolo, T. Heittola, H. Huttunen and T. Virtanen, "Convolutional Recurrent Neural Networks for Polyphonic Sound Event Detection," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 25, no. 6, pp. 1291-1303, June 2017, doi: 10.1109/TASLP.2017.2690575.

[2] D. Stowell, D. Giannoulis, E. Benetos, M. Lagrange and M. D. Plumbley, "Detection and Classification of Acoustic Scenes and Events," in IEEE Transactions on Multimedia, vol. 17, no. 10, pp. 1733-1746, Oct. 2015, doi: 10.1109/TMM.2015.2428998.

[3] H. Phan, M. Maaß, R. Mazur and A. Mertins, "Random Regression Forests for Acoustic Event Detection and Classification," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 23, no. 1, pp. 20-31, Jan. 2015, doi: 10.1109/TASLP.2014.2367814.

[4] G. Valenzise, L. Gerosa, M. Tagliasacchi, F. Antonacci and A. Sarti, "Scream and gunshot detection and localization for audio-surveillance systems," 2007 IEEE Conference on Advanced Video and Signal Based Surveillance, 2007, pp. 21-26, doi: 10.1109/AVSS.2007.4425280.

[5] J. Maxime, X. Alameda-Pineda, L. Girin, and R. Horaud, "Sound representation and classification benchmark for domestic robots," 2014 IEEE International Conference on Robotics and Automation (ICRA), 2014, pp. 6285-6292, doi: 10.1109/ICRA.2014.6907786.

[6] N. Ma, T. May, H. Wierstorf and G. J. Brown, "A machine-hearing system exploiting head movements for binaural sound localisation in reverberant conditions," 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015, pp. 2699-2703, doi: 10.1109/ICASSP.2015.7178461.

[7] J. Chen, Y. Huang, and J. Benesty, Audio Signal Processing for Next Generation Multimedia Communication Systems. Kluwer, 2004, ch. 4-5.

[8] I. Kuzminykh, D. Shevchuk, S. Shiaeles, and B. Ghita, "Audio interval retrieval using convolutional neural networks." Internet of Things, Smart Spaces, and Next Generation Networks and Systems, pp. 229–240., 2020

[9] Y. Zhang and M. Martínez-García, "Machine Hearing for Industrial Fault Diagnosis," 2020 IEEE 16th International Conference on Automation Science and Engineering (CASE), 2020, pp. 849-854, doi: 10.1109/CASE48305.2020.9216787.

[10] A. Mesaros, T. Heittola and A. Klapuri, "Latent semantic analysis in sound event detection," 2011 19th European Signal Processing Conference, 2011, pp. 1307-1311.

[11] T. Heittola, A. Mesaros, T. Virtanen, and M. Gabbouj, "Supervised model training for overlapping sound events based on unsupervised source separation," 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 8677-8681, doi: 10.1109/ICASSP.2013.6639360.

[12] H. Phan, T. N. T. Nguyen, P. Koch, and A. Mertins, "Polyphonic Audio Event Detection: Multi-Label or Multi-Class Multi-Task Classification Problem?," ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2022, pp. 8877-8881, doi: 10.1109/ICASSP43922.2022.9746402.

[13] E. Çakir, G. Parascandolo, T. Heittola, H. Huttunen and T. Virtanen, "Convolutional recurrent neural networks for polyphonic sound event detection", IEEE/ACM Trans. on Audio Speech and Language Processing, vol. 5, no. 6, pp. 1291-1303, 2017

[14] S. Jung, J. Park, and S. Lee, "Polyphonic sound event detection using convolutional bidirectional LSTM and synthetic databased transfer learning", Proc. IEEE Int. Conf. on Acoust. Speech Signal Process, pp. 885-889, 2019.

[15] Y. Fu, K. Xu, H. Mi, H. Wang, D. Wang, and B. Zhu, 'A Mobile Application for Sound Event Detection', in Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19, 7 2019, pp. 6515–6517.

[16] T. Tegou, I. Kalamaras, K. Votis and D. Tzovaras, "A low-cost room-level indoor localization system with easy setup for medical applications," 2018 11th IFIP Wireless and Mobile Networking Conference (WMNC), 2018, pp. 1-7, doi: 10.23919/WMNC.2018.8480912.

[17] D. Čabarkapa, I. Grujić and P. Pavlović, "Comparative analysis of the Bluetooth low-energy indoor positioning systems", Telecommunication in Modern Satellite Cable and Broadcasting Services (TELSIKS) 2015 12th International Conference on, pp. 76-79, 2015.

[18] A. N. Raghavan, H. Ananthapadmanaban, M. S. Sivamurugan, and B. Ravindran, "Accurate mobile robot localization in indoor environments using Bluetooth", Robotics and Automation (ICRA) 2010 IEEE International Conference on, pp. 4391-4396, 2010.

[19] P. K. Yoon, S. Zihajehzadeh, B.-S. Kang and E. J. Park, "Adaptive Kalman filter for indoor localization using Bluetooth low energy and inertial measurement unit", Engineering in Medicine and Biology Society (EMBC) 2015 37th Annual International Conference of the IEEE, pp. 825-828, 2015.

[20] K. J. Piczak, 'ESC: Dataset for Environmental Sound Classification', in Proceedings of the 23rd Annual ACM Conference on Multimedia, 2015, pp. 1015–1018.

 ISBN: ISBNFILL

An Ear Canal Deformation based Head Gesture Recognition Using In-ear Wearables

Youngone Lee
Department of Computer Science
Trinity University
San Antonio, Texas, USA
email: ylee5@trinity.edu

Sheng Tan
Department of Computer Science
Trinity University
San Antonio, Texas, USA
email: stan@trinity.edu

***Abstract*—Hands-free interfaces have become increasingly popular due to the growing demands for convenient control/interaction with mobile and wearable devices. Among all of the hands-free interfaces, head gestures interaction has shown great potential in providing alternatives under various real-world scenarios such as interfaces for people with disabilities, Virtual/Augmented Reality (VR/AR), and vehicle driving. However, existing head gesture recognition systems require either Line-Of-Sight (LOS) or the user to wear specialized hardware. Additionally, those approaches could raise potential privacy concerns. In this work, we propose a novel in-ear wearable system that can achieve head gesture recognition by utilizing off-the-shelf earbuds with a built-in microphone. Specifically, we leverage the relationship between the deformation of the ear canal and the head motion to distinguish different head gestures. A preliminary study shows our system can achieve over 94% recognition accuracy for various head gestures.**

***Keywords*—wearable; human computer interaction; head gesture; Internet-of-Things (IoT).**

I. Introduction

Up until recently, Human Computer Interactions (HCIs) on mobile devices have been dominated by contact interactions including touching the screen or pressing physical buttons. Because of the technological advancement of hardware along with the booming development of ubiquitous computing, a growing number of mobile and wearable devices (e.g., smart glasses, Internet of things devices, virtual reality/augmented reality devices, in-ear wearable devices) have been developed. To better facilitate the control over those emerging devices, more and more hands-free interfaces have been proposed such as gaze tracking [1], voice/speech interaction [2], brain wave control [3], and head posture recognition [4]. Among those novel approaches, head gesture recognition has shown great potential in providing alternatives for various real-life applications. For example, people with certain disabilities and drivers can leverage head gestures to interact with mobile and wearable devices [4]. Additionally, such an approach can be used to control head-mounted VR/AR devices.

Much research effort has been dedicated to developing different techniques for head gesture recognition. Traditionally, Computer Vision (cv) based approaches utilize cameras that can capture the image of the user's head motion to achieve gesture recognition. But, such a solution cannot work under Non-Line-Of-Sight (NLOS) scenarios and suffers from performance degradation in poor lighting conditions. Moreover, CV based approach could raise serious privacy concerns if the image data of the users is not managed properly.

Another body of work leverages motion sensors or Radio Frequency (RF) devices worn by the users to achieve head gesture recognition [4]. The motion sensor-based approach mainly relies on sensors such as accelerometers, and gyroscopes to infer the user's head motion speed and direction. On the other hand, RF devices (e.g., Radio-frequency Identification (RFID) tags, WiFi transceivers) mounted on the users can be used to measure the relative distance to the access point for head gesture recognition. However, those approaches all require specialized or customized hardware, which incurs non-negligible deployment costs. Additionally, some users might be reluctant or feel uncomfortable wearing additional devices.

In this paper, we aim to resolve those issues by proposing an in-ear wearable based head gesture recognition system. This work takes advantage of the Commercial Off-The-Shelf (COTS) earbuds to infer various head movements. It is done by sensing the unique ear canal deformation that closely correlates with distinctive head motions. The proposed system does not require any specialized or additional hardware other than COTS earbuds. Furthermore, our system is unobtrusive to the user during the recognition process and can enhance system security by leveraging user biometrics.

In particular, our system exploits the acoustic sensing approach that can detect the unique ear canal deformation caused by the head motion. The proposed system utilizes a sonar-like technique that can be implemented using any COTS earbud with a built-in microphone. To achieve this, the earbud speaker continuously sends an inaudible acoustic signal through the ear canal when the user is performing a head gesture. The in-ear microphone will capture the signal reflections that encompass ear canal deformation information. Next, our system will analyze the captured signal reflections to detect and distinguish various head gestures. We evaluate our system with a preliminary study and achieve over 94% accuracy in recognizing different head gestures.

The rest of the work is structured as follows. In Section II, we briefly describe the system flow. In Section III, we present the results of the preliminary study. Finally, in Section IV, we

 ISBN: ISBNFILL

conclude and discuss future work.

II. System Design

The underlying principle of our head gesture recognition system lies in the fact that, when a user is performing the head gesture, the motion generated would result in dynamic ear canal deformation with distinctive features. Our system leverages COTS in-ear wearable devices to sense the unique ear canal deformation associated with the head movement of the user for gesture recognition. As shown in Figure 1, our system consists of four major components: *HGI (Head Gesture Interface) Activation*, *Signal Collection*, *Signal Processing*, and *Head Gesture Recognition*. The system first requires HGI Activation to initialize the recognition process which can be triggered by a particular head gesture of the user's choice (e.g., nodding the head or shaking the head). After the activation, the earbud speaker continuously emits an inaudible chirp signal (e.g., over 16kHz) to probe the ear canal. The signal reflected from the ear canal when the user is performing a head gesture will be captured by the inward-facing microphone, which can be further used to extract ear canal deformation information. For the *Signal Processing* component, we first apply various denoising techniques on the collected signal to reduce the inference contained in the captured signal reflections. Next, the denoised signals will go through the segmentation process to find the starting and ending times of the corresponding head gesture. It is done by leveraging the fact that the captured signal reflection will remain relatively consistent when there is no or minimal head motion. Then, the system will move on to the feature extraction component. In particular, we utilize the time-frequency analysis to extract the acoustic characteristics that represent the dynamic ear canal deformation. Lastly, *Head Gesture Recognition* will identify the gesture through the Support Vector Machine (SVM) based classification module and compare it against the pre-built user profile. If the head gesture is identified as unknown, our system will prompt the user to either perform the head gesture again or enroll the unknown gesture into the user profile.

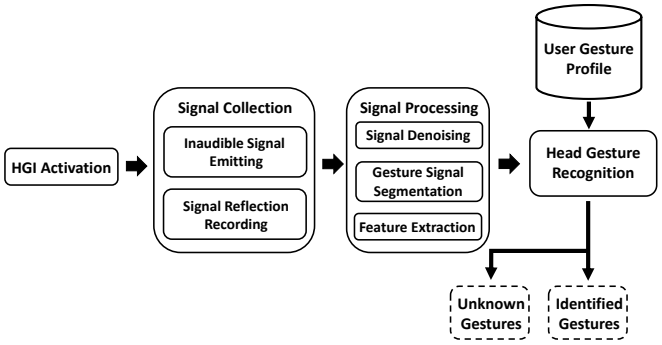


Fig. 1. Overview of system flow.

III. Feasibility Study

To demonstrate the feasibility of the proposed system, we built a prototype device utilizing a COTS in-ear earbud with an embedded microphone chip. The microphone is inward-facing and located in the center area of the speaker. We use Google Pixel 4a with Android 12 that connects to the prototype to control the inaudible probe signal emitting and the reflected signal recording. A chirp signal range from 16kHz to 23kHz is used for the probe signal. We designed four commonly used head gestures inspired by existing work [5]: down and up, up and down, clockwise rotation, and counter-clockwise rotation. Four participants were recruited - two females and two males for the feasibility study. We collected 100 samples from each participant by asking them to perform each head gesture 25 times in their manner. The environments involved in the study are the typical living room and bedroom area. The results are shown in Figure 2. We observe that the proposed system can achieve overall recognition accuracy of around 94%. This demonstrates that our system can effectively recognize various head gestures across different users.

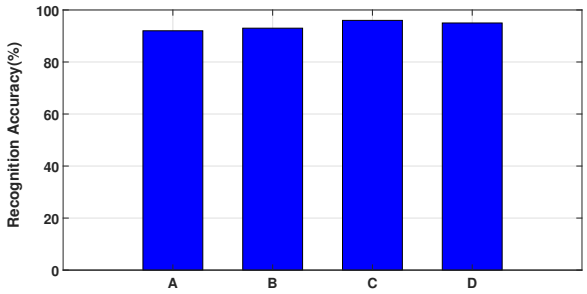


Fig. 2. Recognition accuracy of four different head gestures (A: down and up; B: up and down; C: clockwise rotation; D: counter-clockwise rotation).

IV. Conclusion and Future Work

In this work, we propose a head gesture recognition system utilizing COTS in-ear wearable devices which does not require LOS or any specialized sensor to work. The preliminary study shows that our system can recognize various head gestures with high accuracy. We plan to include more experiments under various scenarios/environments and use more sophisticated deep learning algorithm to further improve recognition accuracy in the future.

References

[1] C. H. Morimoto and M. R. Mimica, "Eye gaze tracking techniques for interactive applications," *Computer vision and image understanding*, vol. 98, no. 1, pp. 4–24, 2005.

[2] C. M. Rebman Jr, M. W. Aiken, and C. G. Cegielski, "Speech recognition in the human–computer interface," *Information & Management*, vol. 40, no. 6, pp. 509–519, 2003.

[3] C. K. Ho and M. Sasaki, "Brain-wave bio potentials based mobile robot control: wavelet-neural network pattern recognition approach," in *2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace (Cat. No. 01CH37236)*, vol. 1, pp. 322–328, IEEE, 2001.

[4] K. Chen, F. Wang, M. Li, B. Liu, H. Chen, and F. Chen, "Headsee: Device-free head gesture recognition with commodity rfid," *Peer-to-Peer Networking and Applications*, vol. 15, no. 3, pp. 1357–1369, 2022.

[5] Y. Yan, C. Yu, X. Yi, and Y. Shi, "Headgesture: Hands-free input approach leveraging head movements for hmd devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 4, pp. 1–23, 2018.

 ISBN: ISBNFILL

Anti-Spoofing for Single-Antenna Devices using Rotational Channel State Information

Avishek Mukherjee*, Tyler Moody†, Mason Strawn‡ and Manish Osti§
Department of Computer Science and Information Systems, Saginaw Valley State University
University Center, Michigan, USA
Email: *amukher1@svsu.edu, †tpmoody@svsu.edu, ‡mwstrawn@svsu.edu, §mrosti@svsu.edu

***Abstract*—In this paper, we investigate the efficacy of using rotational Channel State Information (CSI) on Anti-Spoofing methods in indoor wireless networks. Physical layer information like the CSI often acts like a fingerprint for different locations. Most Anti-Spoofing (AS) methods leverage this uniqueness to detect spoofed packets originating from an attacker that claims to be a genuine user. However, due to the sparsity of wireless channels, there are times when the CSI from different locations may look similar and AS systems may fail to detect a spoofed packet. We propose Rotational Channel State Information - Anti-Spoofing (RCSI-AS), that uses multiple CSI measurements by rotating the antenna on an Access Point (AP) at different angles to greatly improve the detection of spoofed packets. We conducted real world experiments and found that RCSI-AS can detect spoofed packets over 99.6% of the time when using multiple angular configurations at the Access Point (AP) and maintains a low false positive ratio when comparing packets from the same user.**

***Index Terms*—Channel State Information, Wireless Networks, Anti-Spoofing, Wireless Security.**

I. INTRODUCTION

Wireless security has become a critical component of modern Wi-Fi systems with the rapidly increasing production of internet capable devices. As the number of connected devices have increased, so has the potential for different types of network attacks from adversaries. These include channel jamming, packet sniffing, replay attacks, packet spoofing and more. Spoofing, in particular, is a form of attack whereby an adversary (say, Bob) impersonates a user (Alice) device by using their IP address to generate packets to a server. The server may believe these packets originated from Alice and can inadvertently let Bob access confidential information. Detection of spoofing attacks has been a topic of interest for researchers who have proposed anti-spoofing systems to thwart these types of attacks.

In recent years, a number of anti-spoofing methods have been developed that rely on the physical layer information of a wireless packet to determine its authenticity. For example, one of the authors of this paper developed Time-Bounded Anti-Spoofing (TBAS) [1] , that uses physical layer characteristics including the Channel State Information (CSI) to detect spoofed packets. In a wireless system, the CSI between a transmitter and a receiver is represented as a set of complex numbers that is measured at the receiver. This is useful in understanding the channel characteristics and is often used to set rate and beamforming parameters at the sender. In practice, the measured CSI is actually a representation of the multi-path components of the wireless signal from the transmitter to the receiver, due to physical phenomena like scattering, diffraction etc. Thus, in addition to measuring the channel coefficients, the CSI can also be thought of as a fingerprint for a receiver-transmitter location pair. To be more specific, changing the location of either device will result in a different set of measurements as the wireless signal may now undergo a different set of reflections, resulting in different paths. The uniqueness of CSI at different locations can thus be leveraged to identify if a packet arriving at an AP is genuine or spoofed by comparing it with a known CSI measurement from the actual user device.

While the general idea of using CSI measurements to differentiate between users works well, recent studies [2] have also highlighted the sparse nature of wireless channels, whereby the CSI from different locations may exhibit similar patterns making it hard for an Anti-Spoofing system to distinguish between some locations. The idea behind using rotational CSI measurements for spoof detection stems from two interesting observations in industry trends. First, there has been an exponential increase in Internet of Things (IoT) devices in the last decade and these devices now form a large percentage of wireless devices that are constantly communicating with a server. These devices are low power devices and are usually equipped with a single antenna for wireless communications. As such, improvements that rely on antenna diversity may not always be applied to these devices. Second, modern APs such as the Archer AXE200 Omni [3], now come equipped with mechanical antennas that can be rotated automatically with internal motors. These are capable of quickly rotating to different angles and are used to optimize the wireless signal to connected devices.

This poses an interesting question. *Can anti-spoofing methods be improved by using the CSI measurements from different angles on the AP antenna for devices that are limited to a single antenna?* The answer lies in determining whether offsetting the antenna at arbitrary angles introduces sufficient changes in the multi-path components of the signal to be able to distinguish it from the signal at another location. We tackle this problem using experimental analysis in real world locations and propose RCIS-AS - an improved anti-spoofing algorithm complementary to most existing solutions.

The rest of the paper is organized as follows. Section II

 ISBN: ISBNFILL

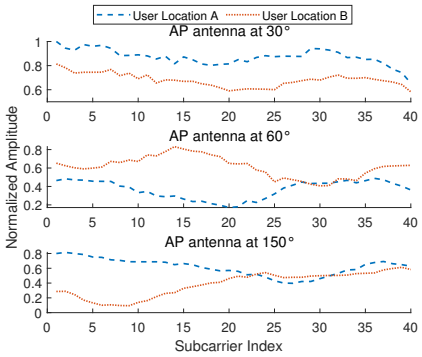
discusses existing research efforts with anti-spoofing. Section III provides a high level overview of RCSI-AS and some theoretical background. Section IV discusses the details of RCSI-AS. Section V evaluates the performance on real world data. Section VI concludes the paper.

II. Related Work

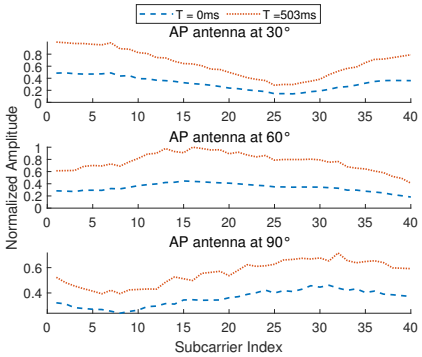
Anti-spoofing methods that rely on physical layer information have been of interest to researchers in recent years. Typically, detection methods that rely on physical layer information usually attempt to localize the spoofed packet using characteristics like the received signal strength [4] [5] or angle or arrival [6] [7] to distinguish between users. RCSI-AS differs from all of these methods as it solely relies on the channel state information for detection of spoofed packets.

There has been some related research that uses the CSI measurements as a fingerprint to detect spoofed packets. These include prior work done by the authors on Time Bounded Anti-Spoofing (TBAS) [1] and Time Bounded Anti-Spoofing on Multiple Input Multiple Output (TBAS-MIMO) systems [8] which is an extension to TBAS with multiple antenna configurations. The key idea behind TBAS is that the CSI for a location does not change in a short interval. When an AP running TBAS receives a packet from a user, it sends a dummy packet to the user that forces the user to send back an ACK in accordance with the 802.11 MAC protocol. If this was a spoofed packet that was sent by an attacker, the AP may receive two responses, one from the actual user (if the attacker does not respond) or a collided signal if both the attacker and the actual user decide to respond. It can then compare the CSI from the original packet and the dummy ACK to determine the authenticity of the request. TBAS was implemented using Software Defined Radios that uses the CSI measurements as well as the power and other physical layer information to achieve low false negative ratios during evaluation. TBAS-MIMO was an experimental study on extending TBAS to commercial off-the-shelf wireless cards that only reports the CSI and no other information. A scenario where the original TBAS sometimes fails is when the CSI from both the attacker and the user coincidentally look similar. This is possible due to the sparsity of wireless channels. TBAS-MIMO attempts to solve this problem by introducing more spatial diversity and comparing the CSI from multiple antenna pairs. TBAS-MIMO also looked at the effects of mobility on the system and recommended guidelines for implementing TBAS. RCSI-AS uses a completely different method from both TBAS and TBAS-MIMO. First, RCSI-AS is targeted towards single antenna devices that may not always be able to take full advantage of the recommendations outlined by TBAS-MIMO. Second, RCSI-AS attempts to introduce spatial diversity by rotating the antenna on the AP to collect CSI measurements at different angular configurations within a short interval. Finally, RCSI-AS periodically probes the user device to frequently update the CSI measurements from the user.

Other research findings that also rely on the CSI include Support Vector Machine (SVM) [9] techniques that uses clustering to distinguish CSI measurements from different users and needs a burn in period. More recent efforts include [10]–[13] which are actually complementary to RCSI-AS. We also note that RCSI-AS is much simpler to implement than some of the methods listed here since RCSI-AS only works at the AP and no modification is needed on the user device. Finally, spoof detection methods like [14] [15] are aimed towards 5G networks with different physical layer characteristics than Wi-Fi.



(a) Different user locations



(b) Same user location 500ms apart

Fig. 1. CSI measurements at different angles

III. Overview of Anti-Spoofing in RCSI-AS

This section provides an overview of RCSI-AS. To keep things simple, we will consider a 1x1 configuration i.e. the AP has a single antenna and the user device also has a single antenna. RCSI-AS works only at the AP and periodically sends out a burst of C probes, where C is referred to as the angular configuration or the number of angles the antenna on the AP is rotated by. The user device receives these probes and responds with ACK packets that are used to measure the CSI between the user and the AP. The value of C is determined empirically, and we found that using $C = 3$ served a good balance between accuracy of RCSI-AS and the probe overhead. As an example, Fig. $1(a)$ shows the CSI measured from two user devices A and B at different locations in a classroom. It can be seen that, when $C = 1$, or when a single probe is used with the antenna positioned at 30 degrees, the CSI measured for the two users look very similar. They seem to be different only by a constant factor which could be attributed to the hardware gain applied to signal during measurement. Thus, using the CSI

 ISBN: ISBNFILL

measured from a single antenna, it may not always be possible to distinguish between two users. However, when using $C = 3$, it is clear that the users are different as the CSI measured at other angles (60 degrees and 150 degrees) look very different from the first configuration. Thus, RCSI-AS would correctly be able to distinguish between user locations in this example. The details on the choice of C can found in Section V.

On the other hand, Fig. $1(b)$ shows the measured CSI on 3 antennas from the same user measured at the AP around 500 milliseconds apart. It is evident that the CSI looks very similar at all 3 angles and the CSI values are only offset by a fixed constant at the AP. This is the key idea behind RCSI-AS, to increase the spatial diversity when antenna diversity is not possible.

In the following sections, we expand on the techniques and heuristics used in RCSI-AS to detect spoofed packets using multiple angular measurements.

IV. Details of RCSI-AS

This section outlines the details of the RCSI-AS system. We note that, while the core idea behind RCSI-AS and its implementation is completely different from our prior work on TBAS, some of the mathematical computations, namely the packet alignment and curve distortion, remain similar as these are metrics used when comparing the similarity of the measured CSI between two packets and can be applicable to any AS system that uses the CSI to detect spoofed packets.

A. Channel State Information

The CSI is measured at the AP and is a set of complex numbers representing the summation of the multiple signal propagation paths from the sender antenna to the receiver antenna. Modern Wi-Fi systems implement Orthogonal Frequency-Division Multiplexing (OFDM). Thus, the CSI measurement on each OFDM subcarrier can be approximated as

$$H = \sum_{p=1}^{P} \alpha_p e^{if\delta_p} \tag{1}$$

where α_p and δ_p denote the amplitude and delay of path p, and P is the total number of multi-path components from the sender to the receiver.

RCSI-AS considers the absolute value for each complex subcarrier when looking at the measured CSI. The phase values of the measured CSI can sometimes comprise of linear and non-linear phase errors [16] that make it difficult to utilize the phase values directly with RCSI-CS.

B. Packet Alignment

As seen in Fig. 1, when measuring the CSI across packets, the hardware applies a different gain value for each packet. Thus, before looking at the differences in CSI between 2 packets, the absolute values need to be aligned. The alignment ratio r between two packets A and B can defined as

$$r = \frac{\sum_{j=1}^{N} a_j b_j}{\sum_{j=1}^{N} a_j^2} \tag{2}$$

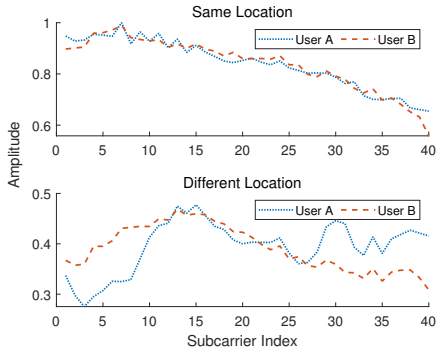


Fig. 2. Effect of alignment on CSI from same location vs different locations

where a_j and b_j are the measured CSI values at subcarrier j for A and B, respectively. The top half of Fig. 2 shows an example of applying the alignment to two packets from the same location that are only offset by some factor. After alignment, both signals almost overlap each other and look very similar in their shape and magnitude. On the other hand, applying the alignment to packets from different locations will still result in very different looking packets. It should also be noted that the alignment ratio is computed for every angular configuration.

C. Curve Distortion

The Curve Distortion ϵ between two aligned packets A and B is defined as

$$\epsilon = \frac{\sum_{j=1}^{N} (ra_j - b_j)^2}{\sum_{j=1}^{N} (ra_j)^2} \tag{3}$$

where a_j and b_j are the aligned CSI values at subcarrier j for A and B, respectively and r is the alignment ratio. This is essentially a numeric representation of the relative difference between two packets. For reference, the ϵ values in Fig. 1(a) when comparing the CSI for different users at $30°, 60°$ and $150°$ are 0.003, 0.17 and 0.35 respectively, whereas for the same user location, as seen in Fig. 1(b), these values are 0.0005, 0.001 and 0.002 It is evident that using multiple angular measurements has a clear advantage as the probability of two distinct user locations exhibiting similar channel characteristics at all of the different angular configurations remains very low.

D. Spoof Detection Threshold

The Curve Distortion is computed for each location pair across every angle in an angular configuration. Suppose we use an angular configuration C of the antenna. We can then define the Spoof Detection Value γ between 2 users X and Y as

$$\gamma = \max_{\{C_k^\circ\}} [\epsilon(k)] \tag{4}$$

where $\epsilon(k)$ refers to the k^{th} curve distortion in C. The ϵ is then compared against a Spoof Detection Threshold to determine if the packets originated from different user

 ISBN: ISBNFILL

locations. We use a threshold of 0.03 in our evaluation of RCSI-AS as we found lower distortion values usually just originate from comparing the Gaussian noise between two similar CSI signals.

V. Evaluation

RCSI-AS was evaluated using real world experimental data collected over the course of one month at different locations in an university setting. The robustness of RCSI-AS was measured using its false positive and false negative performance. The evaluation process is described below.

A. Experimental Setup

As the drivers for routers like the one mentioned in [3] are proprietary, RCSI-AS was built using commercial off-the shelf routers fitted with an external motor and a Raspberry PI to control the rotation of the antenna. An overview of the hardware setup and architecture is shown in Fig. 3. The AP used in RCSI-AS is a TP-Link N750 Wireless router. The router's firmware was modified with the OpenWrt [17] tool that enabled greater control of the channel, rate, and transmission power parameters. In addition, one of the external antennas of the router was connected to a Dorhea SG90 Micro Servo Motor that was controlled using a Raspberry Pi 3 Model B. To measure the CSI, we installed the NexMon CSITool [18] [19] on a Raspberry PI Model 4 which contains a single internal antenna. An external laptop was used to send pings to the router that enabled us to measure the CSI. Some auxiliary switches were used to facilitate remote operation of these devices.

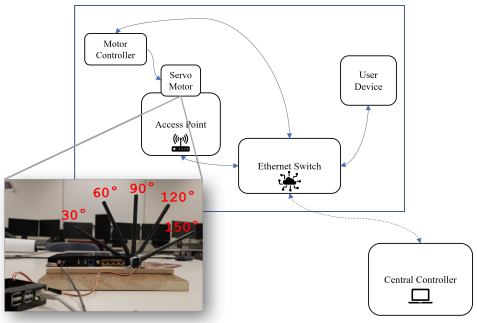


Fig. 3. System Architecture

B. Data Collection

We ran several experiments using the Nexmon CSITool at different locations including classrooms, computing labs and office spaces. We enabled only a single antenna for the AP and so the measured CSI was a linear vector representing the 64 subcarrier values on a 20MHz wireless channel. The experiments were conducted in an environment with relatively moderate mobility. This allowed us to simulate different types CSI data representative of real world wireless channels. Some example locations can be seen in Fig. 4 where the AP was kept stationary inside a computer lab while the CSI for different user locations was measured. For each user location, a total of 5 angular CSI measurements were recorded. These measurements were recorded at $[30°, 60°, 90°, 120°, 150°]$. We note that these angles were chosen arbitrarily. In other words, RCSI-AS will work with any configuration of angular measurements. Our evaluation suite consisted of measurements from 500 different locations.

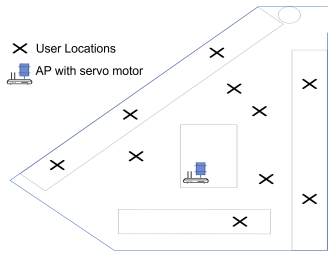


Fig. 4. Some Experimental Locations

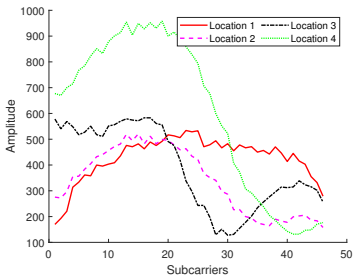


Fig. 5. Absolute value of CSI in 4 locations

C. Data Preprocessing

The CSI measurements were first sanitized by removing null and pilot subcarrier indexes, as defined in the 802.11ac standard [20]. These subcarriers do not contain actual measurements and are not considered by RCSI-AS. The absolute value of some typical CSI measurements after removal is shown in Fig. 5. In addition, a few more pre-processing operations were performed.

- It can be seen from Fig. 5 that the subcarriers at both ends of the measurement seem to attenuate. The exact cause of this is unknown, although we suspect it is due to some additional filtering in hardware. Thus, we truncate the signal by removing 6 subcarriers from both ends of the measured CSI.
- Sometimes the recorded CSI contains very high spikes that do not represent actual measurements. We discard measurements where the number of subcarriers with spikes exceed 30% of the total number of measured subcarriers.
- In some cases, the CSI was only recorded for half of the bandwidth. These packets were also filtered out by comparing the average power between the first and second halves of the CSI measurement and filtering these packets out if the difference between them exceeded an order of magnitude.
- After performing the above preprocessing steps, we observed that the measured CSI may still contain some

 ISBN: ISBNFILL

outlier values. Thus a final round of pre-processing is performed using a Hampel filter to detect these outliers from the CSI measurements using a median absolute deviation threshold over a window size of 5. We note that this does not always smooth out all the outliers, especially at either end of the measured CSI, but this is a limitation of the CSITool not RCSI-AS.

- After pre-processing the measured CSI was normalized with its maximum amplitude set to 1.

D. False Negative Performance

This section describes the false negative performance of RCSI-AS. The goal of this evaluation is to check how often RCSI-AS can correctly identify that two users X and Y are different by comparing their measured CSI. The preprocessed CSI measurements from all 500 locations was considered for this evaluation. A random CSI measurement was chosen from each location. Then, the comparison described in Section IV was performed for each unique location pair.

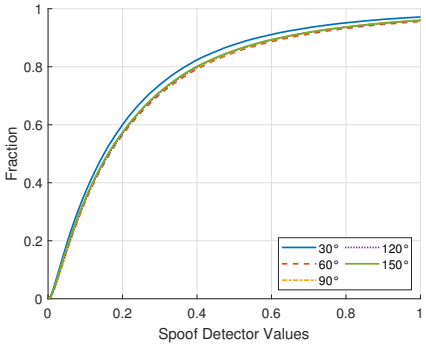


Fig. 6. False Negative Performance with Single Angular Configuration

1) Base Case: Fig. 6 shows the cumulative distribution of spoof detector values (γ) when using angular configuration of $C = 1$. It can be clearly seen that in the base case, when only one angle of the antenna is considered, roughly 8.98% of the CSI pairs may be misclassified by RCSI-AS as the same user using a spoof detection threshold of 0.03. A typical example was shown earlier in Fig. 1(a) where the measured CSI from two different locations looked similar. Upon further inspection, Fig. 7 shows a distribution of the location indexes along with the fraction of misclassifications by RCSI-AS when considering only the measured CSI at one angle. It can be seen that for most locations, there is at least one other location where the measured CSI may coincidentally look similar.

2) Multiple Angular Configurations: The advantage of RCSI-AS becomes evident when moving to higher angular configurations ($C > 2$), which reduces the probability that the measured CSI will exhibit similar characteristics across all angles in C. We looked at the performance of RCSI-AS for every combination of different angular configurations. For example, when looking at angular configurations of size 2, the performance of all $\binom{5}{2}$ or 10 possible combinations was considered and it was found that the percentage of misclassifications drops to 1.1% as opposed to almost 9% in the

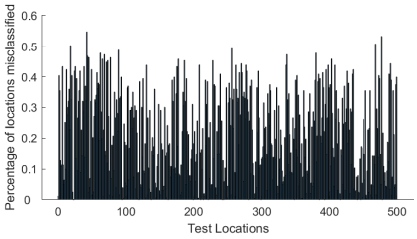


Fig. 7. Percentage of Locations that are misclassified by RCSI-AS at 90°

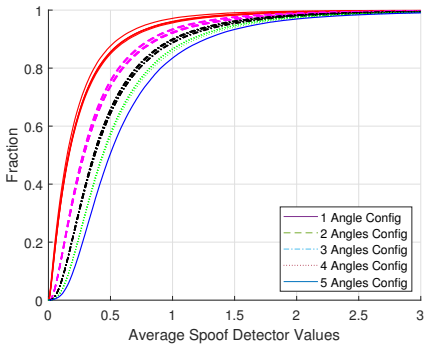


Fig. 8. False Negative Performance with higher configurations

base case. Fig. 8 shows the cumulative distribution plot of all possible angular configurations. It is clear that increasing the angular measurements to RCSI-AS results in a much lower mis-classification rate. When all antenna angles are used $C = 5$, the mis-classification rate drops to 0.16% which makes RCSI-AS extremely accurate, albeit at the expense of a higher probe overhead. Based on the empirical data, we found that using $C = 3$ or 3 angular measurements serves as a good balance between the probe overhead and results in overall accuracy of 99.69%.

E. False Positive Performance

This section discusses the misclassification rate of RCSI-AS when considering packets from the same user. To measure its performance, we looked at packets from the same user location at different intervals during the data collection process. RCSI-AS compared a total of over 4800 packet pairs from same locations measured within a short interval of each other.

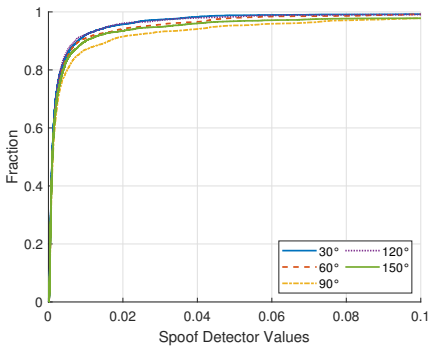


Fig. 9. False Positive Performance with Single Angular Configuration

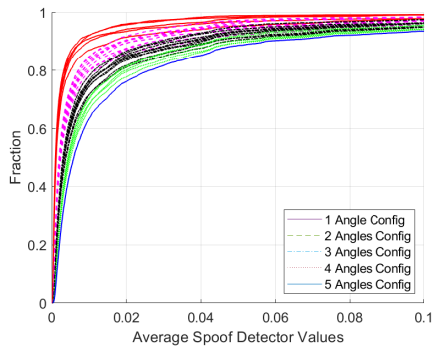


Fig. 10. False Positive Performance with higher configurations

1) Base Case: As with the false negative evaluation, we first establish the performance RCSI-AS when using $C = 1$. Fig. 9 shows the cumulative distribution of the spoof detector values (γ) across each individual angle when comparing the CSI from a single angular measurement. It can be seen that the false positive ratio is very good, and the percentage of misclassifications is below 5% when using a threshold of 0.03.

2) Higher Configurations: It can be seen from Fig. 10 that increasing the number of angular configurations has a slight degradation on the false positive performance. This is expected since we consider the largest curve distortion value within each combination. The false positive performance still remains around 93% when using an angular configuration of $C = 3$ values. In addition, Fig. 11 shows the distribution of the error values on all subcarriers when comparing the signals after alignment at every angle. The distribution is mostly smooth and the differences mainly arise from the quantization and noise in the measured CSI.

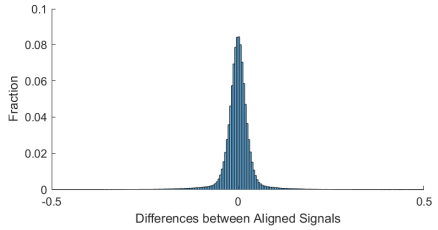


Fig. 11. False Positive Performance when using higher configurations

VI. Conclusion and Future Work

We proposed RCSI-AS, a novel anti-spoofing system based on rotational channel state information on commodity wireless APs. We implement a motorized system that allows rotation of the antennas of an AP to measure the CSI at multiple angles. RCSI-AS works by periodically sending multiple probes to a user device at different angular configurations. This allows RCSI-AS to detect spoofed packets even when the CSI at different locations may exhibit similar characteristics at one angular position of the antenna. We evaluated RCSI-AS using real world experiments in 500 different locations and found that RCSI-AS can detect packets from different locations over 99.60% of the time when using 3 or more angular configurations. At the same time, when looking at packets from the same user location RCSI-AS will correctly identify packets from the same user over 95% of the time. RCSI-AS is aimed at single antenna devices which may not always be able to take advantage of anti-spoofing methods that rely on antenna diversity. We are currently looking into extensions to RCSI-AS that can reduce the probe overhead even further and also assess its performance in high mobility environments.

References

[1] M. Liu, A. Mukherjee, Z. Zhang, and X. Liu, "TBAS: Enhancing Wi-Fi Authentication by Actively Eliciting Channel State Information," in *2016 13th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pp. 1–9, 2016.

[2] R. He, B. Ai, G. Wang, M. Yang, C. Huang, and Z. Zhong, "Wireless Channel Sparsity: Measurement, Analysis, and Exploitation in Estimation," *IEEE Wireless Communications*, vol. 28, no. 4, pp. 113–119, 2021.

[3] "AXE11000 Tri-Band Wi-Fi 6E Router." https://www.tp-link.com/us/home-networking/wifi-router/archer-axe200-omni/, 2013.

[4] Y. Chen, W. Trappe, and R. P. Martin, "Detecting and Localizing Wireless Spoofing Attacks," in *2007 4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, pp. 193–202, 2007.

[5] J. Yang, Y. Chen, W. Trappe, and J. Cheng, "Detection and localization of multiple spoofing attackers in wireless networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 1, pp. 44–58, 2013.

[6] H.-C. Chen, T.-H. Lin, H. T. Kung, C.-K. Lin, and Y. Gwon, "Determining RF angle of arrival using COTS antenna arrays: A field evaluation," in *MILCOM 2012 - 2012 IEEE Military Communications Conference*, pp. 1–6, 2012.

[7] J. Xiong and K. Jamieson, "SecureAngle: Improving Wireless Security Using Angle-of-Arrival Information," Association for Computing Machinery, 2010.

[8] A. Mukherjee, A. W. Garvin, S. E. Sanchez, and Z. Zhang, "Experimental Evaluation of Time Bounded Anti-Spoofing (TBAS) in MIMO Systems," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pp. 1–6, 2017.

[9] H. Liu, Y. Wang, J. Liu, J. Yang, and Y. Chen, "Practical User Authentication Leveraging Channel State Information (CSI)," Association for Computing Machinery, 2014.

[10] J. K. Tugnait, "Detection of Pilot Spoofing Attack Over Frequency Selective Channels," in *2018 IEEE Statistical Signal Processing Workshop (SSP)*, pp. 737–741, 2018.

[11] C. Li and A. Sezgin, "Spoofing attack detection in dynamic channels with imperfect CSI," *arXiv preprint arXiv:2101.06185*, 2021.

[12] X. Li, K. Huang, S. Wang, and X. Xu, "A physical layer authentication mechanism for IoT devices," vol. 19, pp. 129–140, 2022.

[13] Z. Jiang, J. Zhao, X.-Y. Li, J. Han, and W. Xi, "Rejecting the attack: Source authentication for Wi-Fi management frames using CSI Information," in *2013 Proceedings IEEE INFOCOM*, pp. 2544–2552, 2013.

[14] W. Li, N. Wang, L. Jiao, and K. Zeng, "Physical Layer Spoofing Attack Detection in MmWave Massive MIMO 5G Networks," vol. 9, pp. 60419–60432, 2021.

[15] D. Spoljar, K. Lenac, D. Zigman, and M. Marović, "A Mobile Network-Based GNSS Anti-Spoofing," in *2018 26th Telecommunications Forum (TELFOR)*, pp. 1–3, 2018.

[16] H. Zhu, Y. Zhuo, Q. Liu, and S. Chang, "Pi-splicer: Perceiving accurate csi phases with commodity wifi devices," *IEEE Transactions on Mobile Computing*, vol. 17, no. 9, pp. 2155–2165, 2018.

[17] "OpenWrt: A Linux kernel based operating system for embedded solutions." https://openwrt.org/, 2022.

[18] M. Schulz, D. Wegemer, and M. Hollick, "Nexmon: The C-based Firmware Patching Framework." https://openwrt.org/, 2017.

[19] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free Your CSI: A Channel State Information Extraction Platform For Modern Wi-Fi Chipsets," in *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization*, WiNTECH '19, p. 21–28, 2019.

[20] "IEEE Standard for Information technology – Telecommunications and information exchange between systems," *IEEE Std 802.11ac-2013*, pp. 1–425, 2013.

 ISBN: ISBNFILL

Secure-by-Design Methodology using Meet-in-the-Middle Design Flow for Hardware Implementations of ECC-based Passive RFID Tags

Manh-Hiep Dao*†, Vincent Beroulle*, Yann Kieffer*, Xuan-Tu Tran†
*Univ. Grenoble Alpes, Grenoble INP, LCIS, 26000 Valence, France
†VNU Information Technology Institute, Vietnam National University, Hanoi, Vietnam
Corresponding author's email: tutx@vnu.edu.vn

***Abstract*—With the rapid development of needs concerning the secured passive Radio Frequency IDentification (RFID) tag, several works propose the implementation of authentication protocols based on Elliptic Curve Cryptography (ECC). But, there are no systematical approaches that allow considering both the compatibility of security and the implementation cost as part of the requirements and limitations of passive RFID tags. Therefore, the problem of balancing the implementation cost and the security requirements for passive RFID tags is still an open question. In this paper, we present a part of a Security-by-Design methodology that targets passive RFID tags using ECC primitives.**

***Index Terms*—Passive RFID, Side-Channel Attack, Design Methodology, Security by Design, ECC, Meet-in-the-Middle.**

I. Introduction

Passive RFID tags are portable devices utilizing radio frequency to authenticate the identity of the objects. By communicating via a wireless channel, these devices face a variety of vulnerabilities such as wireless and hardware attacks. The wireless threats try to illegally access the system to steal or modify the data communicated via the wireless channel. Hardware attacks, such as Side-Channel Attacks (SCA) and Fault Attacks (FA), exploit the weaknesses of design to reveal the secret key. In order to protect the secret information contained in the passive RFID tag, the device tends to implement a secured authentication protocol using cryptography primitives.

Among the cryptography primitives, the asymmetric encryption algorithms, which are based on the Discrete Logarithm Problem (DLP), are recommended to replace the symmetric ones to avoid the key distribution issue. In the asymmetric family, Elliptic Curve Cryptography (ECC) is more attractive due to the advance of shortened key length. However, compared to other cryptography primitives such as symmetric encryption algorithm (AES [1], PRESENT [2]) or hash function, ECC as well as asymmetric cryptography methods are much more complicated. That leads to higher implementation costs such as physical area, power consumption, and latency. Meanwhile, the design of passive RFID tags is very challenging because of the constraints of implementation. Therefore, looking for a compatible design methodology that balances the protection of security and identity data privacy with minimizing the implementation cost of passive RFID tags is still an open question for researchers.

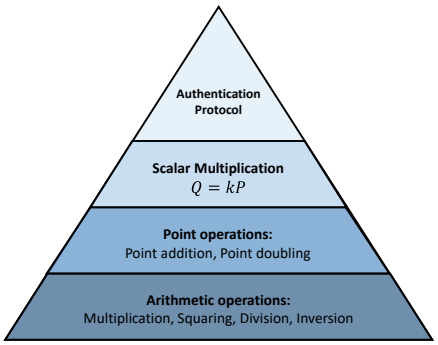


Fig. 1. Concept of Authentication Protocol using ECC primitives.

In the literature, the secured authentication protocols that use cryptography primitives, for passive RFID tags are implemented based on the concept including four primary abstraction levels, as seen in Fig. 1. There are already several design flows that allow taking security into account. The most popular approach used for designing ECC primitives is the top-down design methodology, which is applied in several works [3]–[8] in literature. However, because they lack information on the systematical architecture, these implementations can not prove the compatibility of their design with the design constraints of passive RFID tags.

In addition, there are some works [9], [10] that apply the Bottom-Up Design Methodology. One of the disadvantages of the bottom-up design methodology is that it is time-consuming. When the combined system is too complicated, the simulation and verification become much more complex. Because of this issue, this design methodology is not compatible with designing ECC primitives. In both design methodologies mentioned above, they misconstrue the impact of security as an additional feature of the design. Therefore, it is difficult to obtain an ECC-based authentication protocol balancing implementation cost and security requirements.

In order to solve the mentioned problem, we propose to use a Security-by-Design methodology based on the classical Meet-in-the-Middle approach for designers. A combination of the recommended design methodology with our proposal

 ISBN: ISBNFILL

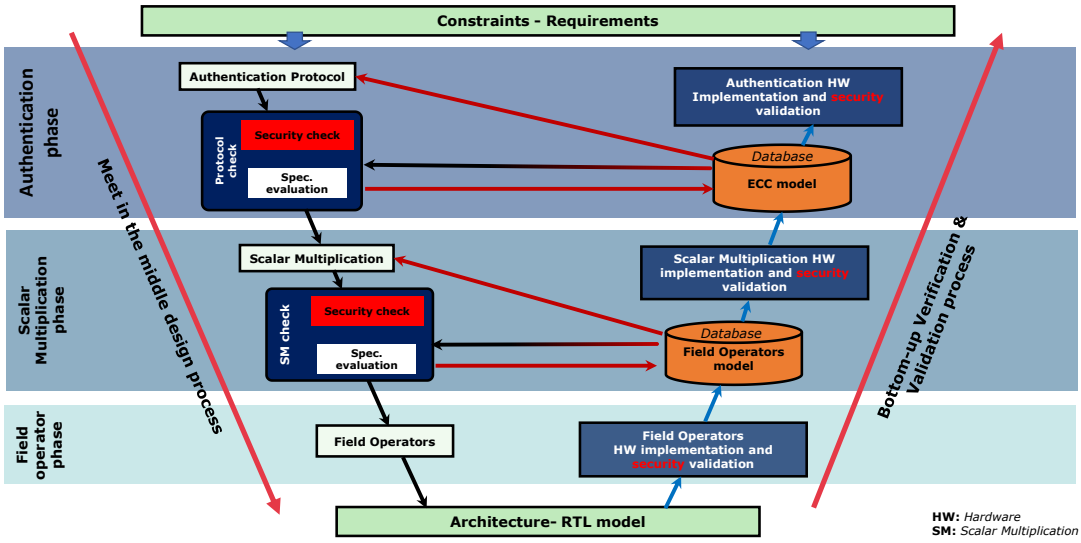


Fig. 2. Proposed Security-by-Design Methodology using Meet-in-the-Middle.

helps designers to consider both hardware security issues and the implementation cost of authentication protocol using ECC primitive. Consequently, the final authentication protocol based on ECC primitives balances both the security requirements and the design constraints of passive RFID tags. Specifically, our contributions in this paper are:

- A proposed Security-by-Design methodology based on the classical Meet-in-the-Middle approach.
- An estimation of implementation cost and SCA vulnerabilities assessment method for the ECC block based on the knowledge of field operator primitives in the literature.

The organization of this paper is as follows. Section II presents our proposal. Section III indicates an example of the proposed design methodology. In the last section, a conclusion gives a summary of the work and the remaining tasks in the future.

II. Security by Design Methodology

In this section, we present our proposed Security-by-Design Methodology using Meet-in-the-Middle, as described in Fig. 2. This strategy combines the Meet-in-the-Middle Design and the Bottom-Up Evaluation and Validation process. At the beginning of the design process, designers determine the specifications of the target system in terms of implementation cost and security requirements by analyzing choices with the knowledge of the sub-blocks. The final result of this design process is finding a configuration and architecture for the ECC primitive with a compatible authentication protocol.

In the following, the Bottom-Up Evaluation and Validation process is carried out after implementing and assembling the sub-blocks into the system. The aim of this process is to assess the security and validate the design by implementing them on hardware.

A. Meet-in-the-Middle Design Process

This process comprises three phases: Authentication, Scalar Multiplication, and Field Operators Design phase, as described in Fig. 2. Due to the page limitation of this paper, we only discuss in more detail the last two phases of the Meet-in-the-Middle Design Process.

1) Authentication Phase: In the Authentication Phase, according to the knowledge of the ECC primitive blocks, which are proposed in the literature, designers consider both the security and implementation cost of various protocols. After choosing a compatible authentication protocol, designers would know the design constraints and security requirements of the ECC primitive at the beginning of the second phase.

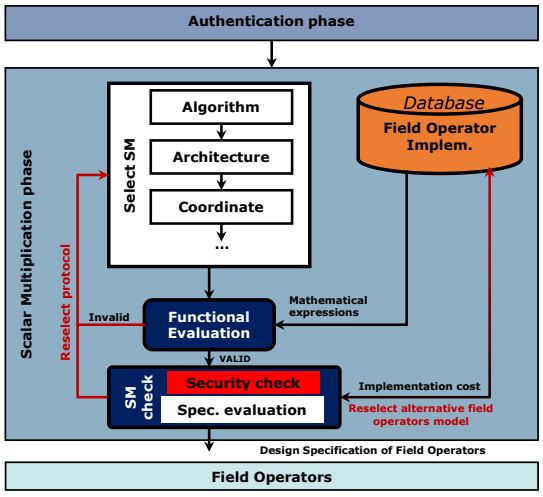


Fig. 3. Scalar Multiplication Design Phase.

2) Scalar Multiplication Phase: In this phase, firstly, based on the determined specification of the ECC primitive, a process of choosing a field, algorithm, architecture, and projective coordinates system would be done, as depicted in Fig. 3.

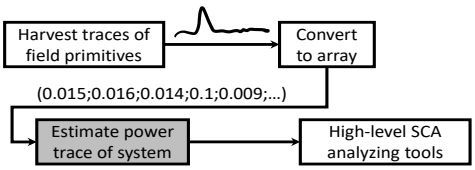


Fig. 4. Process of estimating power trace of the ECC primitive.

In the first sub-step, behavioral and abstraction models are extracted from the previous work in the literature. Behavioral models express the mathematical formulas of the field operators, for example, field multiplier, inversion, and squaring. They are used by designers to verify the functionalities of Scalar Multiplication. The abstraction model of field operators indicates the configuration, implementation cost, and power traces with corresponding data. The designers use these abstraction models for evaluating the security and estimating the implementation cost in the Evaluation and Estimation step.

a) Implementation cost estimation of the ECC block: The parameters that we can estimate are the area (A_{total}), power consumption (P_{total}), and maximum duration (T_{total}) of the ECC design. By knowing the detailed configuration of the system, designers could know how many field operators are needed to carry out scalar multiplication. In the following, we note N_i the number of required field operators with the index of $\{1; 2; 3\}$ being field multiplier, field square, and field inverter, respectively. We also note L_{key} the number of bits of the key.

Depending on the architecture of the ECC system, designers know about the layout of sub-blocks and the interconnection of the target system. Therefore, we can estimate the area of the ECC block by using Eq. 1. We note n_i the number of sub-blocks implementing field operators, A_i, as the area of each sub-block carrying out the field operator i.

$$A_{total} = \sum_{i=1}^{3} n_i \cdot A_i \tag{1}$$

Besides, the maximum duration and power consumption of the ECC system are estimated via Eq.2 if the field operators compute in parallel. We note (P_i, T_i) the power consumption and delay of each sub-block carrying out the field operator, respectively i.

$$\begin{aligned} P_{total} &= \sum_{i=1}^{3} n_i \cdot P_i \\ T_{total} &= L_{key} \cdot \max\{(\frac{N_i}{n_i} \cdot T_i) : i = (1, 2, 3)\} \end{aligned} \tag{2}$$

If these sub-blocks work sequentially, the power and maximum duration of the ECC system are estimated via Eq.3.

$$\begin{aligned} P_{total} &= \max\{P_i : i = (1, 2, 3)\} \\ T_{total} &= L_{key} \cdot \sum_{i=1}^{3} \frac{N_i}{n_i} \cdot T_i \end{aligned} \tag{3}$$

b) Power traces estimation: In more detail, our proposal shows an approach to assess approximated power traces of scalar multiplication by the reference traces of primitive operators. At the beginning of this step, we collect the power traces of the field operators, as described in Fig.4 together with corresponding data. These traces could be harvested from the silicon device or estimated by using the power simulator tools. After collecting the traces, they are converted into arrays $P_i = \{p_{i0}, p_1, \cdots, p_{ik}\}$. We note p_{ij} the simultaneous power at the moment $t = j$ of operator i, and k is the length of the trace. In the next step, these arrays are used for estimating the trace of the ECC system.

Depending on the architecture of the ECC primitives, these arrays could be concatenated or accumulated to form the power trace of the system. In the case, two operators, which have power traces P_1, P_2, compute in parallel, the total power trace is the accumulation of P_1 and P_2 as Eq.4:

$$P_{total} = \sum_{i=0}^{k} (P_{1i} + P_{2i}) \tag{4}$$

On the opposite, if two operators work sequentially, the total power trace is the concatenation of two sub-traces as Eq.5. Consequently, the length of the total power trace is longer.

$$P_{total} = \{P_1 | P2\} = \{p_{10}, p_{11}, \cdots, p_{1k}, p_{20}, p_{21}, \cdots, p_{2k}\} \tag{5}$$

At the end of this step, the power trace of the system will be assessed by the high-level pre-silicon SCA evaluation tools such as CASCADE [11], or ChipWhisperer [12].

3) Field Operators Phase: In the last phase of the top-down design process, there is a specification of field operators such as field multiplier, square, and perhaps divider as the result of the scalar multiplication phase. The final result of this phase is implementing a full architecture of field operators before interconnecting them to obtain the system of the ECC primitive design. At the end of this phase, designers implement the chosen field operators in the RTL (Register-Transfer Level) model, and then, begin the bottom-up evaluation and validation process.

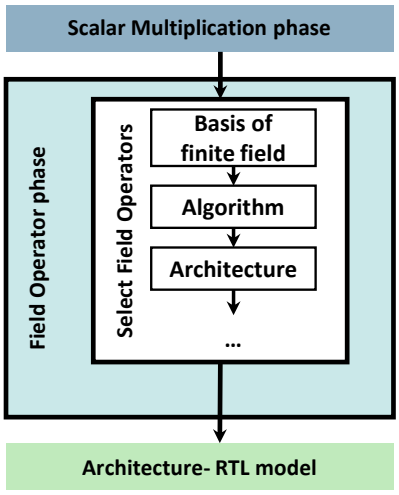


Fig. 5. Field Operators Design Phase.

B. Bottom-up Verification and Validation Process

The inputs of the bottom-up evaluation process are the final RTL model of the full architecture of the ECC-based authentication protocol with the reference models of primitive operators. In each bottom-up evaluation phase, there are two steps: security evaluation and estimation of the cost of input. There are 3 main evaluation phases as discussed below.

1) Field Operator Evaluation phase: The first evaluation phase is Field Operator, as illustrated in Fig. 6. After implementing the RTL model of field operators, by using the EDA (Electronic Design Automation) synthesis tools, designers verify their functionalities. In addition, both estimating the implementation cost and collecting the real power traces of field operators are also carried out on hardware platforms.

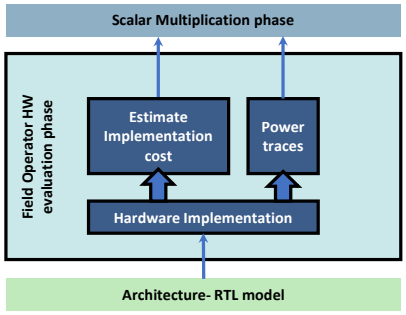


Fig. 6. Field Operators Evaluation Phase.

2) Scalar Multiplication Evaluation Phase: After evaluating and validating the field operators, these sub-blocks are assembled into the ECC system carrying out the scalar multiplication. The input of this evaluation phase is the hardware implementation of the ECC system. At the beginning of the Scalar Multiplication Evaluation Phase, as demonstrated in Fig.7, the power traces of scalar multiplication are provided to the post-silicon evaluation tool for SCA assessment with the corresponding input data. If the evaluation is failed, designers have to go back to re-select the algorithm of Scalar Multiplication or countermeasures for ECC.

After the successful SCA evaluation of hardware, designers also measure the implementation costs of the ECC primitives. If the implemented ECC primitive is overpriced compared to the reference model, designers go back to the beginning and choose another architecture and algorithm for the Scalar Multiplication block. Conversely, they update the new optimized ECC primitives on the database and embed the ECC primitive to the chosen authentication protocol before continuing to implement the final design of the ECC-based authentication protocol.

III. Example of Meet-in-the-Middle Design Process

In this section, an example of our proposed Secure-by-Design Methodology is presented. Specifically, the process of selecting a design for an ECC-based authentication protocol

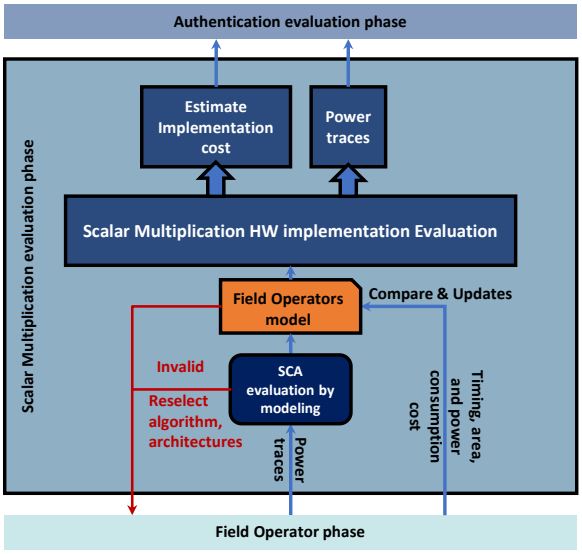


Fig. 7. Scalar Multiplication Evaluation Phase.

implementation with design constraints and security requirements is described. The final design specification satisfies all the design constraints and security requirements of the passive RFID tag, which uses ASIC as an implementation platform.

The assumption constraints of the ECC-based authentication protocol are listed below:

- Implementation costs:
  - Maximum duration of one tag-server authentication: T_{auth} = 20 ms
  - Maximum available power: P_{max} = 240 μW
  - Implementation area: as low as possible
- Security properties:
  - Secure against Simple SCA and Differential SCA.

The maximum timing of communication refers to the standard ISO/IEC-14443, meanwhile, the maximum power consumption is the peak harvesting at $-3dBm$ incident power by rectifier antenna, according to the proposal in [13].

A. Authentication Phase

We follow the analysis of Gabsi [14] and choose Zhao's protocol [15]. Gabsi showed that this protocol is secure against the Differential SCA. In Zhao's protocol, the tag performs five scalar multiplications. Thus, the maximum duration for one scalar multiplication is $4ms$ $(= 20ms/5)$. Regarding the security requirements, since the protocol is already secure against differential SCA, the scalar multiplication implementation will only have to be secure against Simple SCA.

B. Scalar Multiplication Phase

After determining the target specifications of scalar multiplication, we start the Scalar Multiplication phase, as declared in Fig. 3.

1) Select field: Wenger et al. [20] recommended choosing the Binary Field $GF(2^m)$. In the hardware implementation, Wenger showed that this field requires less than Prime Field

TABLE I
IMPLEMENTATION COST OF DIFFERENT BINARY ELLIPTIC CURVES IN PROJECTIVE COORDINATE.

Curve	Coordinate System	Cost of Point Multiplication	Complete
Binary Generic Curve [16]	Mixed	$6M + 4S$	×
Binary Edward Curve [17]	Mixed	$6M + 4S$	✓
Binary Edward Curve [18]	Affine	$I + 11M + 4S$	✓
Binary Edward Curve [18]	Projective	$16M + S$	✓
Generalized Hessian Curve [19]	Mixed	$9D + 4S$	✓

* I, M, S denote the field inverter, multiplier, and square, respectively.

$GF(p)$ in the context of implementation cost. In addition, we follow standard FIPS 186-4 and choose the 163-bit length of the key for the passive RFID tag.

2) Select algorithm: During this step, we continue to select the configuration of the curve and the algorithm of scalar multiplication. Firstly, there are several curves that can be implemented in the $GF(2^m)$. Table I lists different curves in $GF(2^m)$. Based on Table I, we chose the Binary Edward Curve in the mixed coordinate to implement the point multiplication. This curve has a completeness property that makes it robust against the Simple SCA and also Differential Side-Channel Attacks [21]. In addition, this curve requires $6M+4S$ less than other choices.

TABLE II
EXAMPLE OF THE INITIALIZATION PHASE: FIELD OPERATORS.

	Montgomery multiplier	Montgomery square
Area (kGates)	3.3	0.6
Power (μW)	63	51
Latency (μs)	8.5	8.5

3) Choosing the reference field operators: Before evaluating and estimating the implementation cost of the scalar multiplication block, we choose the reference field operators. As the analysis above, our design only needs field multipliers and field squares. Therefore, we take the implementation cost of these operators which is presented by Deschamps et al. [22]. The synthesis results that we obtained with the ASIC technology NangateOpencore $45\,\mathrm{nm}$ with the maximum clock frequency of $20\,\mathrm{MHz}$ are shown in Table II.

4) Select architecture: In this section, we only consider 3 different architectures with different levels of parallelism of the multipliers and the squares. Arch. a) comprises 3 sub-blocks of field multiplier and 2 sub-blocks of field square. Arch. b) and c) includes 2 and 1 sub-blocks of field multiplier and field square, respectively.

5) Implementation cost estimation: In this step, the implementation costs of the three previous architectures are computed via Equations 1-3. The results of parallel computing in field multiplier and field square are given in Table III.

TABLE III
IMPLEMENTATION COST ESTIMATION FOR SCALAR MULTIPLICATION USING PARALLEL COMPUTING.

Archi.	Area (kGates)	Power (μW)	Latency (ms)
Arch. a)	11.1	291	2.7
Arch. b)	7.8	228	4.13
Arch. c)	3.9	114	8.07

In Table III, we choose Arch.b) for implementing the scalar multiplication as it satisfies the maximum available power and requires an acceptable area. Although its latency is much close to our expectation ($4ms$), it will be improved by optimizing the field operators. The specification of the maximum duration for optimizing field operators in the next phase is determined by Eq. 6.

$$\frac{4}{2 \times 163} \approx 8.18(\mu s) \tag{6}$$

6) Power traces Estimation: Because using 163-bit of key length, there are 163 loop iterations in one scalar multiplication. In each loop, based on the chosen architecture, 2 sub-blocks of field multiplier and 2 sub-blocks of field square are parallel computing. Therefore, based on Eq. 4, the total power traces of scalar multiplication is the accumulation of power traces of sub-blocks.

In the third step, we also estimate the power trace of each field operator by the power estimator tool in $8.5(\mu s)$, which is the duration of an operation. Power traces of 1^{st} and 2^{nd} field multiplier sub-blocks are indicated as the orange dot line and green dash line in Fig. 8. Red dot-dash and purple lines in Fig. 8 illustrate the power traces of 1^{st} and 2^{nd} field square sub-blocks. The total power trace is described as the blue line in Fig. 8. This trace is assessed by the high-level pre-silicon SCA tools.

At the end of this design phase, by performing 6 steps of choosing, evaluating, and validating, we find the target configuration of scalar multiplication as below:

- Field: Binary field $GF(2^{163})$
- Curve: Binary Edward Curve in Mixed Coordinate
- Architecture: Using 2 sub-blocks of field multiplier and 2 sub-blocks of field square
- Parallel computing in field multiplier and field square

Besides, we also determine the specification of the field operators. They need to optimize to achieve 8.18 (μs) of latency. Furthermore, in this step, we also estimate the power trace of the system for assessment by pre-silicon SCA evaluation tools.

IV. CONCLUSION AND FUTURE WORK

In this paper, we present a part of the Secure-by-Design methodology using the Meet-in-the-Middle approach that enables the designer to obtain an ECC primitive block that balances the security level and the implementation costs.

 ISBN: ISBNFILL

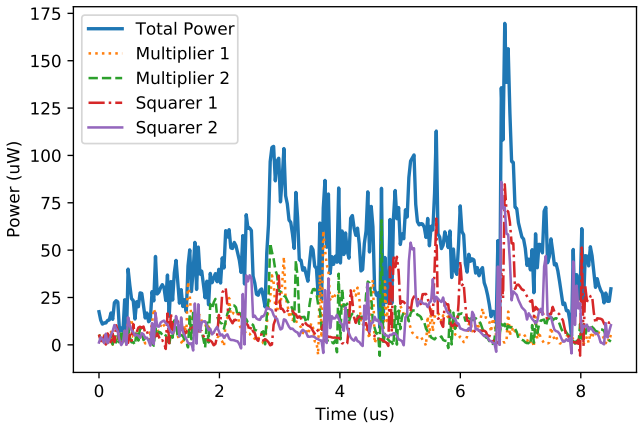


Fig. 8. Example of Power Traces Estimation.

Furthermore, our proposed estimation method enables an approximation of the implementation costs and also the security level of the ECC primitive block based on the knowledge of previous field operator primitives in the literature. In the future, we will include the selection and evaluation of the authentication protocols in the Secure-by-Design Methodology.

ACKNOWLEDGMENT

This work is supported by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02) and Vietnam National University, Hanoi under project QG-ECC-RFID, code QG.22.70. The first author is partly funded by Vingroup JSC and supported by the Master, Ph.D. Scholarship Programme of Vingroup Innovation Foundation (VINIF), Institute of Big Data, code VINIF.2021.TS.166

REFERENCES

[1] M.-H. Dao, V.-P. Hoang, V.-L. Dao, and X.-T. Tran, "An energy efficient aes encryption core for hardware security implementation in iot systems," in *2018 ATC*. IEEE, 2018, pp. 301–304.

[2] A. Bogdanov *et al.*, "Present: An ultra-lightweight block cipher," in *Cryptographic Hardware and Embedded Systems-CHES 2007: 9th International Workshop, Vienna, Austria, September 10-13, 2007. Proceedings 9*. Springer, 2007, pp. 450–466.

[3] R. Salarifard, S. Bayat-Sarmadi, and H. Mosanaei-Boorani, "A low-latency and low-complexity point-multiplication in ecc," *IEEE TCAS-I: Regular Papers*, vol. 65, 2018.

[4] T. Shahroodi, S. Bayat-Sarmadi, and H. Mosanaei-Boorani, "Low-latency double point multiplication architecture using differential addition chain over gf(2m)," *IEEE TCAS-I: Regular Papers*, vol. 66, 2019.

[5] R. Salarifard and S. Bayat-Sarmadi, "An efficient low-latency point-multiplication over curve25519," *IEEE TCAS-I: Regular Papers*, vol. 66, 2019.

[6] P. Choi, M. K. Lee, J. H. Kim, and D. K. Kim, "Low-complexity elliptic curve cryptography processor based on configurable partial modular reduction over nist prime fields," *IEEE TCAS-II: Express Briefs*, vol. 65, 2018.

[7] S. R. Pillutla and L. Boppana, "Low-complexity bit-serial sequential polynomial basis finite field gf(2m) montgomery multipliers," *Microprocessors and Microsystems*, vol. 84, 2021.

[8] Q. Shao *et al.*, "Low complexity implementation of unified systolic multipliers for nist pentanomials and trinomials over *GF(2^m)*," *IEEE TCAS-I: Regular Papers*, vol. 65, no. 8, pp. 2455–2465, 2018.

[9] N. Pirotte, J. Vliegen, L. Batina, and N. Mentens, "Balancing elliptic curve coprocessors from bottom to top," *Microprocessors and Microsystems*, vol. 71, p. 102866, 2019.

[10] J. Lutz and M. Anwarul Hasan, "High performance elliptic curve cryptographic co-processor," in *Wireless Network Security*. Springer, 2007, pp. 3–42.

[11] D. Sijacic, J. Balasch, B. Yang, S. Ghosh, and I. Verbauwhede, "Towards efficient and automated side channel evaluations at design time," *Kalpa Publications in Computing*, vol. 7, pp. 16–31, 2018.

[12] C. O'flynn and Z. D. Chen, "Chipwhisperer: An open-source platform for hardware embedded security research," in *COSADE*. Springer, 2014, pp. 243–260.

[13] P. Xu, D. Flandre, and D. Bol, "Analysis, modeling, and design of a 2.45-ghz rf energy harvester for swipt iot smart sensors," *IEEE JSSC*, vol. 54, 2019.

[14] S. Gabsi, V. Beroulle, Y. Kieffer, H. M. Dao, Y. Kortli, and B. Hamdi, "Survey: Vulnerability analysis of low-cost ecc-based rfid protocols against wireless and side-channel attacks," *Sensors*, vol. 21, no. 17, p. 5824, 2021.

[15] Z. Zhao, "A secure rfid authentication protocol for healthcare environments using elliptic curve cryptosystem," *Journal of medical systems*, vol. 38, pp. 1–7, 2014.

[16] J. López and R. Dahab, "Fast multiplication on elliptic curves over *GF*(2^m) without precomputation," in *CHES*. Springer, 1999, pp. 316–327.

[17] B. Koziel, R. Azarderakhsh, and M. Mozaffari-Kermani, "Low-resource and fast binary edwards curves cryptography," in *INDOCRYPT*. Springer, 2015, pp. 347–369.

[18] K. H. Kim, C. O. Lee, and C. Negre, "Binary edwards curves revisited," in *INDOCRYPT*. Springer, 2014, pp. 393–408.

[19] R. R. Farashahi and M. Joye, "Efficient arithmetic on hessian curves," in *PKC*. Springer, 2010, pp. 243–260.

[20] E. Wenger and J. Grossschadl, "An 8-bit avr-based elliptic curve cryptographic risc processor for the internet of things," in *MICRO*. IEEE, 2012, pp. 39–46.

[21] H. M. Edwards, "A normal form for elliptic curves," *Bulletin of the American Mathematical Society*, vol. 44, 2007.

[22] J. P. Deschamps, G. D. Sutter, and E. Cantó, "Guide to fpga implementation of arithmetic functions," *Lecture Notes in Electrical Engineering*, vol. 149 LNEE, 2012.

 ISBN: ISBNFILL