Project Title: Spoof Detection for Automatic Speaker Verification

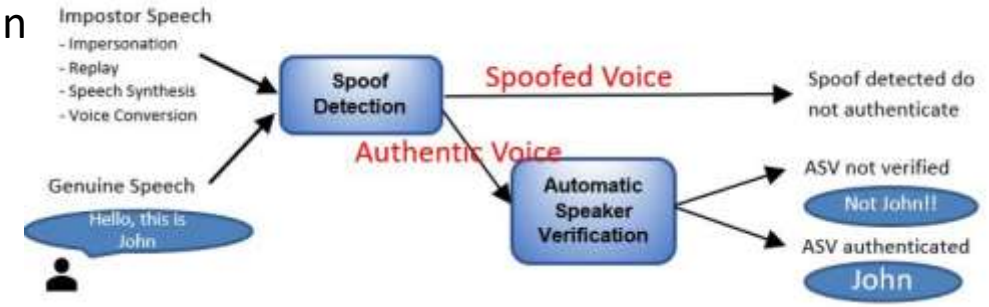
Background :

The human voice is one of the biometric indicators representing a person's identity. It has been successfully used in automatic speaker verification (ASV) systems in many applications. The ASV system is currently vulnerable to spoofing attacks in which someone disguises as another and illegitimately accesses a secure system. Hence, countermeasures against spoofing attacks are necessary to verify whether the claimed voice is a genuine or fake representation before verifying or identifying the speaker's identity. Therefore, this research focuses on spoofing detection in automatic speaker verification.

Targets:

The objectives of this project are listed as follows:

(1) To explore and investigate significant of speech features for spoof detection,
(2) To optimize percentage of voice and non-voice segments in features used in spoofing detection.
(3) To investigate pathological feature for spoof detection
(4) To minimize detection error,
(5) To improve an accuracy of ASV, and
(6) To study multi-lingual spoof detection.



Speaker:

Kasorn GALAJIT(Project Leader and Speaker)
National Electronics and Computer Technology Center, Thailand

Project Title : Spoof Detection for Automatic Speaker Verification

Project Members : 35 members (5 countries, 7 institutes)

Name / Position / Institution	Name / Position / Institution	Name / Position / Institution	Name / Position / Institution	Name / Position / Institution
Masashi Unoki (PhD) / Professor / JAIST, Japan	Waree Kongprawechnon (PhD) / Assoc. Prof. / SIIT, Thailand	Nanthayod Termkoh / Undergrad student / SIIT, Thailand	Win Lai Lai Phyu / Asst. Lecturer / UCSY, Myanmar	Kasorn Galajit (PhD) / Project Leader / NECTEC, Thailand
Candy Olivia Mawalin (PhD) / Asst. Prof. / JAIST, Japan	Pakinee Aimmanee / Assoc. Prof. / SIIT, Thailand	Kosin Kalarat / PhD student / SIIT, Thailand	Myat Aye Aye Aung / Asst. Lecturer / UCSY, Myanmar	Suradej Duangpummet (PhD) / Researcher / NECTEC, Thailand
Kai Li / PhD / JAIST, Japan	Sasiporn Usanavasin(PhD) / Assoc. Prof. / SIIT, Thailand	Pantarat Vichathai/ Undergrad student / SIIT, Thailand	Win Pa Pa (PhD) / Professor / UCSY, Myanmar	Jessada Karnjana (PhD) / Researcher / NECTEC, Thailand
Anuwat Chaiwongyen / PhD / JAIST, Japan & SIIT, Thailand	Khaing Zar Mon / Master student / SIIT, Thailand	Puchit Bunpleng/ Undergrad student / SIIT, Thailand	Aye Mya Hlaing (PhD) / Lecturer / UCSY, Myanmar	Surasak Boonkla (PhD) / Researcher / NECTEC, Thailand
Khalid ZAMAN/ PhD Candidate / JAIST, Japan	Patthranit Kaewcharuay / Undergrad student / SIIT, Thailand	Patharapol Laolakkana/ Undergrad student / SIIT, Thailand	Hay Mar Soe Naing/ Lecturer/ UCSY, Myanmar	Phondanai Khanti / Assistant Researcher / NECTEC, Thailand
Aulia Adila / Master Student / JAIST, Japan	Ananda Garin Mills / Undergrad student/ SIIT, Thailand	Dk Hayati Pg Hj Mohd Yassin (PhD) / Lecturer / UBD, Brunei	Dessi Puji Lestari (PhD) / Professor / ITB, Indonesia	Xugang Lu (PhD) / Senior Researcher / NICT, Japan
Nopparut Li / Master Student / JAIST, Japan	Pannathorn Sathirasattayanon / Undergrad student / SIIT, Thailand	Navod Neranjan Thilakarathne / PhD Candidate / UBD, Brunei	Sarah Azka Arief/ Master student/ ITB, Indonesia	Sheng Li (PhD) / Senior Researcher / NICT, Japan

Project Duration : 2 years, 1 April 2023 – 31 March 2025

Remark : Widhyakorn Asdornwised (PhD) / Asst. Prof. / CU, Thailand is removed

Project Budget: 80,000 USD (40,000 USD per year)

Project Activities: Scientific

- NECTEC, JAIST, SIIT, UBD make the improvement of "ThaiSpoof" Dataset for spoof detection for Thai language
- Incorporating new spoofed voice data generated using the multilingual speech (MMS) model. Researchers can use this dataset to study spoof detection
- SIIT, JAIST, NECTEC Study "Potential of Speech-pathological Features for Deepfake Speech Detection"
  Investigate the significant of pathological feature for deepfake detection
- UBD, NECTEC, JAIST, study Exploring a Cutting-Edge Convolutional Neural Network for Speech Emotion Recognition.
  This study explorer the emotion from speech signal..

Project Activities: Scientific

- UCSY, NECTEC, JAIST create UCSYSpoof: A Myanmar Language Dataset for Voice Spoofing Detection
  The study gives knowledge of voice spoofing in Myanmar language
  Also Pathological studies for Myanmar language

- JAIST, NECTEC, SIIT study Deepfake-speech detection with pathological features and multilayer perceptron neural network
  This study can show the significant of pathological for spoof detection

- ITB JAIST study Anomalous Machine Sound Detection Based on Time Domain Gammatone Spectrogram Feature and IDNN Model
- This study explorer the feature of the sound in time domain to find the anomalous machine sound detection.

Project Activities: Budget plan in detail

Budget: 40,000 USD/year

Meeting/workshop (2 times a year)	14,000 USD
IVO Forum	1,000 USD
Publication, presentation at international conferences	25,000 USD

Meeting (14,000 USD)

May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
Kick-off					2nd	Forum					
3rd						Forum	Final			END	

Kick-off meeting **2,500**USD (Thailand May 2023) ,
Second meeting **6,000**USD (Indonesia Dec 2023)
Third meeting **12,600**USD (Japan Oct 2024)

Publication (First year)

ISAI 2 papers	800USD
SCDE 1 Paper	2,700USD
Cocosda 2 papers	2,1260USD
ACM 1 paper	310USD
APSIPA2023 1 paper	0 USD
IEEE Access 1 Journal	0 USD

Total **5970** USD

Publication (2nd year)

ISAI Thailand 2 papers	400USD
ICNLSP 1 Paper	350USD
Cocosda 3 papers	2,020USD
APSIPA2025	2,000USD
M3 conference	5,000USD
IEEE Access 1 Journal	2,000 USD

Total **11,770** USD

Approximately used budget

- First year Total **15,400** USD
  - Meeting 8,500
  - Publication 5,970
  - Forum 1,000
- 2nd year Total **25,370** USD
  - Meeting 12,600
  - Publication 11,770
  - Forum 1,000

R&D results:

- “ThaiSpoof ”: Dataset for spoof detection for Thai language

  This dataset is published on AI for Thai for sharing
- “UCSYSpoof”: Dataset for spoof detection for Myanmar language
- Dataset for spoof detection for Indonesian language

- Ongoing

  Webpage for spoof detection.
  ASEAN language dataset

Scientific Contribution:

Presentations at International Conferences:

Expected output in 2 years

10 Conference papers

No:	Paper title:	Author names	Affiliation	Conference name:	The date of the conference	The venue of the conference
1	ThaiSpoof: A Database for Spoof Detection in Thai Language	• Kasorn Galajit, • Thunpisit Kosolsriwiwat, • Candy Olivia Mawalim, • Pakinee Aimmanee, • Waree Kongprawechnon, • Win Pa Pa, • Anuwat Chaiwongyen, • Teeradaj Racharak, • Hayati Yassin, • Jessada Karnjana, • Surasak Boonkla, • Masashi Unoki,	• NECTEC, National Science and Technology Development Agency, • Sirindhorn International Institute of Technology, • Japan Advanced Institute of Science and Technology, • University of Computer Studies,Yangon, • Universiti Brunei Darussalam}, Brunei Darussalam	The 18th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP 2023) (https://isai-nlp-aiot2023.aiat.or.th/)	27-29/11/2023	Bangkok, Thailand
2	Spoof Detection using Voice Contribution on LFCC features and ResNet-34	• Khaing Zar Mon, • Kasorn Galajit, • Candy Olivia Mawalim, • Jessada Karnjana, • Tsuyoshi Isshiki, • Pakinee Aimmanee,	• Sirindhorn International Institute of Technology, • NECTEC, National Science and Technology Development Agency, • Japan Advanced Institute of Science and Technology, • Tokyo Institute of Technology Tokyo,	The 18th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP 2023) (https://isai-nlp-aiot2023.aiat.or.th/)	27-29/11/2023	Bangkok, Thailand
3	Using Novel Hybrid Convolutional Neural Network for Dysarthria Diagnosis	• Navod Neranjan Thilakarathne, • Kasorn Galajit, • Jessada Karnjana, • Win Pa Pa, • Candy Olivia Mawali • Hayati Yassin,	• Universiti Brunei Darussalam}, Brunei Darussalam • NECTEC, National Science and Technology Development Agency, • Japan Advanced Institute of Science and Technology, • University of Computer Studies,Yangon,	The 10th IEEE CSDE 2023, the Asia-Pacific Conference on Computer Science and Data Engineering 2023,(IEEE CSDE 2023) https://ieee-csde.org/csde2023/	4-6/12/2023	Yanuca Island, Fiji

Scientific Contribution:

Presentations at International Conferences:

No:	Paper title:	Author names	Affiliation	Conference name:	The date of the conference	The venue of the conference
4	Speech Watermarking for Tampering Detection Using Singular Spectrum Analysis with a Psychoacoustic Model	• Phondanai Khanti, • Pannathorn Sathirasattayanon, • Patthranit Kaewcharuay, • Nanthayod Termkoh, • Ekachai Phaisangittisagul, • Kasorn Galajit, • Jessada Karnjana	• Sirindhorn International Institute of Technology, • NECTEC, National Science and Technology Development Agency, • Kasetsart University	The 26th Conference of the Oriental COCOSDA https://www.ococosda2023.com/	4-6/12/2023	Delhi, India
5	A Large Vocabulary End-to-End Myanmar Automatic Speech Recognition	• Hay Mar Soe Naing • Win Pa Pa	• University of Computer Studies,Yangon,	M3Oriental Workshop of ACM Multimedia Asia 2023 The ACM Multimedia Asia 2023 https://sites.google.com/view/m3oriental	8/12/2023	Tainan city, Taiwan
6	Deepfake-speech detection with pathological features and multilayer perceptron neural network	• Anuwat Chaiwongyen, • Suradej Duangpummet, • Jessada Karnjana, • Waree Kongprawechnon, • Masashi Unoki	• Japan Advanced Institute of Science and Technology, • NECTEC, National Science and Technology Development Agency, • Sirindhorn International Institute of Technology,	The 15th annual conference organized by Asia-Pacific Signal and Information Processing Association (APSIPA2023) https://www.apsipa2023.org/	31/11/2023-3/12/2023	Teipei, Taiwan

Scientific Contribution:

Presentations at International Conferences:

No:	Paper title:	Author names	Affiliation	Conference name:	The date of the conference	The venue of the conference
7	M-Diarization: A Myanmar Speaker Diarization using Multi-scale dynamic weights	Myat Aye Aye Aung, Win Pa Pa, Hay Mar Soe Naing.	• University of Computer Studies,Yangon,,	The 26th Conference of the Oriental COCOSDA https://www.ococosda2023.com/	4-6/12/2023	Delhi, India
8	Exploring a Cutting-Edge Convolutional Neural Network for Speech Emotion Recognition.	Thilakarathne, Navod Neranjan, Kasorn Galajit, Candy Olivia Mawalim, and Hayati Yassin.	• Universiti Brunei Darussalam, Brunei Darussalam, • NECTEC, National Science and Technology Development Agency, • Japan Advanced Institute of Science and Technology	The 2024 5th International Conference on Industrial Engineering and Artificial Intelligence (IEAI)	24-26/04/2024	Bangkok, Thailand
9	UCSYSpoof: A Myanmar Language Dataset for Voice Spoofing Detection	Win Pa Pa, Hay Mar Soe Naing, Aye Mya Hlaing, Kasorn Galajit, Candy Olivia Mawalim	• University of Computer Studies,Yangon, • NECTEC, National Science and Technology Development Agency, • Japan Advanced Institute of Science and Technology	The 27th International Conference of the Oriental COCOSDA 2024	17-19/10/2024	HainChu, Taiwan

Scientific Contribution:

Presentations at International Conferences:

No:	Paper title:	Author names	Affiliation	Conference name:	The date of the conference	The venue of the conference
10	Analysis of Pathological Features for Spoof Detection	Myat Aye Aye Aung,Win Pa Pa, Hay Mar Soe Naing, Aye Mya Hlaing, Kasorn Galajit, Candy Olivia Mawalim	• University of Computer Studies,Yangon, • NECTEC, National Science and Technology Development Agency, • Japan Advanced Institute of Science and Technology	The 27th International Conference of the Oriental COCOSDA 2024	17-19/10/2024	HainChu, Taiwan
11	Speech Watermarking for Tampering Detection using Singular Spectrum Analysis with Quantization Index Modulation and Psychoacoustic Model	Pantarat Vichathai,Puchit Bunpleng, Patharapol Laolakkana, Dr. Sasiporn Usanavasin, Phondanai Khanti, Dr. Kasorn Galajit, Dr. Jessada Karnjana	• Sirindhorn International Institute of Technology, • NECTEC, National Science and Technology Development Agency,	The 27th International Conference of the Oriental COCOSDA 2024	17-19/10/2024	HainChu, Taiwan
12	Generative Adversarial Network based Neural Vocoder for Myanmar End-to-End Speech Synthesis	Aye Mya Hlaing and Win Pa Pa	• University of Computer Studies, Yangon, Myanmar	7th International Conference on Natural Language and Speech Processing https://www.icnlsp.org/2024welcome/	19-20/10/2024	Trento, Italy

Scientific Contribution:

Presentations at International Conferences:

No:	Paper title:	Author names	Affiliation	Conference name:	The date of the conference	The venue of the conference
13	Parallel and Limited Data Voice Conversions on Myanmar Language Speech for Spoofed Detection	Hay Mar Soe Naing, Win Pa Pa, Sheng Li	• University of Computer Studies,Yangon, • National Institute of Information and Communications Technology,	The 2nd Workshop of Multimodal, Multilingual and Multitask Modeling Technologies for Oriental Languages (M3Oriental) https://sites.google.com/view/m3oriental	3/12/2024	Auckland, New Zealand
14	Automatic Speaker Verification on Myanmar Spoofing Voice Data using GMM-UBM and TDNN	Win Pa Pa, Win Lai Lai Phyu, Hay Mar Soe Naing	• University of Computer Studies,Yangon,	The 2nd Workshop of Multimodal, Multilingual and Multitask Modeling Technologies for Oriental Languages (M3Oriental) https://sites.google.com/view/m3oriental	3/12/2024	Auckland, New Zealand
15	Anomalous Machine Sound Detection Based on Time Domain Gammatone Spectrogram Feature and IDNN Model	Dessi Puji Lestari,Primanda Adyatma Hafiz, Candy Olivia Mawalim, Sakriani Sakti, and Masashi Unoki	• Bandung Institute of Technology, Bandung, Indonesia • Japan Advanced Institute of Science and Technology	The 16th APSIPA Annual Summit and Conference (ASC)2024 http://www.apsipa2024.org/index.html	3-6/12/2024	Macau, China

Scientific Contribution: 3 Published Journal Papers

Published Journal Papers:

Expected output in 2 years
10 Conference papers
2 Journals

Current output (approximately 1.5 year)
15 Conference papers
3 Journals

No:	Paper title:	Author names	Affiliation	Journal name:	The publisher of the Journal	The volume number and Pages
1	Contributions of Jitter and Shimmer in the Voice for Fake Audio Detection	KAI LI, XUGANG LU, MASATO AKAGI, MASASHI UNOKI.	• Japan Advanced Institute of Science and Technology, • Advanced Speech Technology Laboratory, National Institute of Information and Communications Technology,	IEEE Access	IEEE Access	VOLUME 11, 2023 Received 21 June 2023, accepted 29 July 2023, date of publication 3 August 2023, Digital Object Identifier 10.1109/ACCESS.2023.3301616
2	Title: End-to-End Neural Diarization for Unknown Number of Speakers with Multi-Scale Decoder	Aung, Myat Aye Aye, Win Pa Pa, Hay Mar Soe Naing	• University of Computer Studies,Yangon	International Journal of Intelligent Engineering & Systems 17, no. 5 (2024).	The Intelligent Networks and Systems Society	Vol 17, Issue 5, p870 ISSN: 2185-310X DOI: 10.22266/ijies2024.1031.66

Scientific Contribution:

Published Journal Papers:

Expected output in 2 years

10 Conference papers
2 Journals

Current output (approximately 1.5 year)

15 Conference papers
3 Journals

No:	Paper title:	Author names	Affiliation	Journal name:	The publisher of the Journal	The volume number and Pages
3	Potential of Speech-pathological Features for Deepfake Speech Detection	ANUWAT CHAIWONGYEN, SURADEJ DUANGPUMMET, JESSADA KARNJANA, WAREE KONGPRAWECHNON and MASASHI UNOKI	• Japan Advanced Institute of Science and Technology, • Sirindhorn International Institute of Technology, • NECTEC, National Science and Technology Development Agency,	IEEE Access	IEEE Access	Received 20 June 2024, accepted 15 August 2024, date of publication 22 August 2024, date of current version 10 September 2024. Digital Object Identifier 10.1109/ACCESS.2024.3447582

Additional Journal papers

- Zaman, Khalid, Melike Sah, Cem Direkoglu, and Masashi Unoki. "A survey of audio classification using deep learning." *IEEE Access* (2023).
- Ota, Yasuji, and Masashi Unoki. "Anomalous sound detection for industrial machines using acoustical features related to timbral metrics." *IEEE Access* (2023).

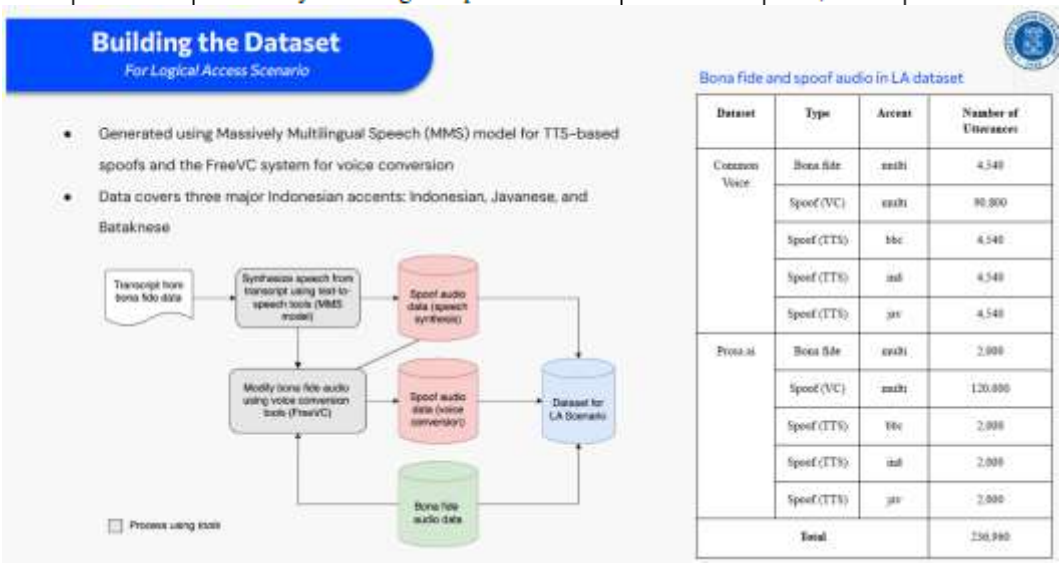
Societal Impact: Dataset

- **ThaiSpoof: Spoof Detection in Thai Language**

Label	Dataset Type	Degree	Utterances
Genuine	Genuine Dataset	-	4,583
Spoof	Text to Speech - TTS	-	4,583
	F0 Modification	10 ch/oct	4,583
		40 ch/oct	4,583
		160 ch/oct	4,583
		320 ch/oct	4,583
	Pitch Shifting	+ 4%	4,583
		+ 10%	4,583
		+ 20%	4,583
		-4%	4,583
		-10%	4,583
		- 20%	4,583
	Massively Multilingual Speech - MMS	-	4,583

- **UCSYSpoof: Spoof Detection in Myanmar Language**

Label	Subset Type	No. of Speaker	No. of Utterance
Genuine	Genuine	3 (4K utts/each)	12,000
Spoofed	Vocoder-based [3]	3	23,932
	FreeVC-based	3	24,000
	Text-to-speech [3]	1	8,000
	GMM VC	2	8,000
	GMM DIFFVC	2	8,000



Bona fide and spoof audio in LA dataset

Dataset	Type	Accent	Number of Utterances
Common Voice	Bona fide	multi	4,540
	Spoof (VC)	multi	90,800
	Spoof (TTS)	bbc	4,540
	Spoof (TTS)	sul	4,540
	Spoof (TTS)	jav	4,540
Proso.ai	Bona fide	multi	2,000
	Spoof (VC)	multi	120,000
	Spoof (TTS)	bbc	2,000
	Spoof (TTS)	sul	2,000
	Spoof (TTS)	jav	2,000
Total			230,980

ENIT DATASET: ENGLISH NATIVE AND INDONESIAN-THAI NON-NATIVE SPEECH

Type	Sets	# unique spk.		# utterances		# spoof attack
		Male	Female	Bonafide	Spoof	
Native	Train.	56	56	5,590	33,540	8
	Dev.	8	8	800	4,800	8
	Eval.	16	16	1,600	9,600	8
Non-native	Train.	10 / 12	4 / 5	5,696	33,164	3
	Dev.	1 / 1	1 / 2	542	4,275	3
	Eval.	3 / 3	2 / 3	1,752	10,501	3

Data Splitting

Train-dev-eval
- 70-10-20
- No overlap speaker

Conclusion:

- Dataset :“ThaiSpoof” Dataset for spoof detection for Thai language
- Myanmar dataset for speaker recognition and diarization.
- The knowledge of the contributions of Jitter and Shimmer in the Voice for Fake Audio Detection
- The knowledge of the contributions of voice and non-voice for detect spoofing in MFCC and LFCC
- The knowledge of characteristic of abnormal voice.
- The knowledge of speech recognition in Myanmar language
- The Knowledge of the significant of pathological for spoof detection
- The Knowledge of characteristic of tampering speech which can be used for spoof detection

Future works:

- “ThaiSpoof ”: Dataset for spoof detection for Thai language

  This dataset is published on AI for Thai for sharing
- “UCSYSpoof”: Dataset for spoof detection for Myanmar language
- Dataset for spoof detection for Indonesian language
- Ongoing

  Webpage for spoof detection.

  ASEAN language dataset

From the Third meeting, members agree to extend project for 1 years. :Prepare application for extension