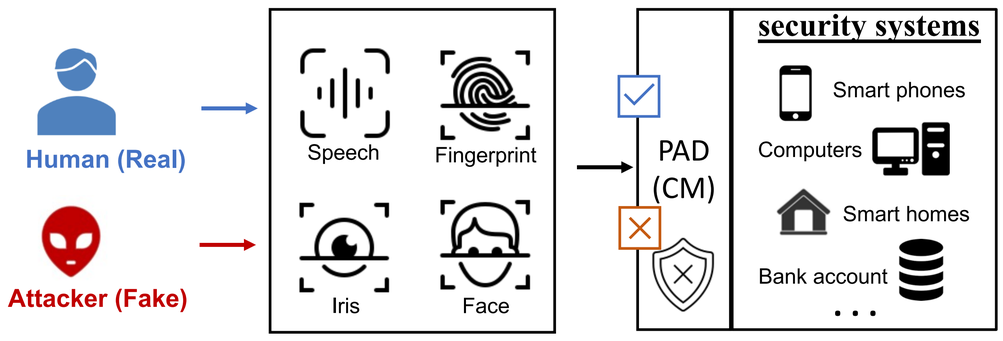
Range-Based Equal Error Rate for Spoof Localization

Lin Zhang[1,2], Xin Wang[1], Erica Cooper[1], Nicholas Evans[3], Junichi Yamagishi[1,2]
[1]National Institute of Informatics, Japan [2]SOKENDAI, Japan [3]EURECOM, France

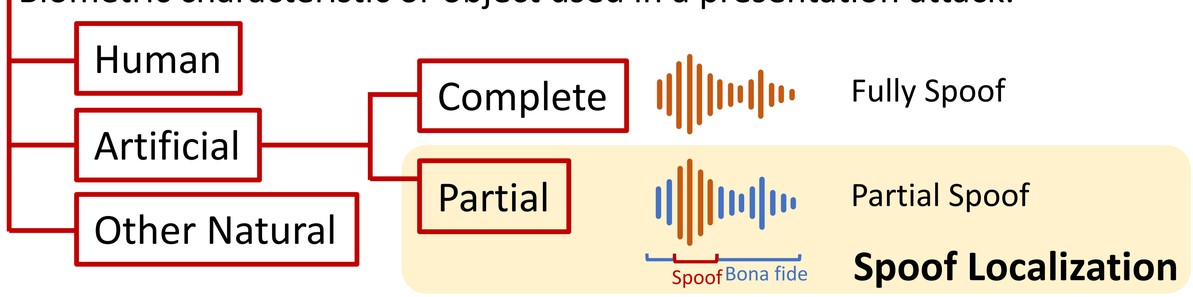
Introduction

ISO/IEC JTC1 SC37 Biometrics: ISO/IEC 30107: Information technology — Biometric presentation attack detection (2016)

Biometric technologies are vulnerable to **Presentation Attack Instrument (PAI)**.



Biometric characteristic or object used in a presentation attack.



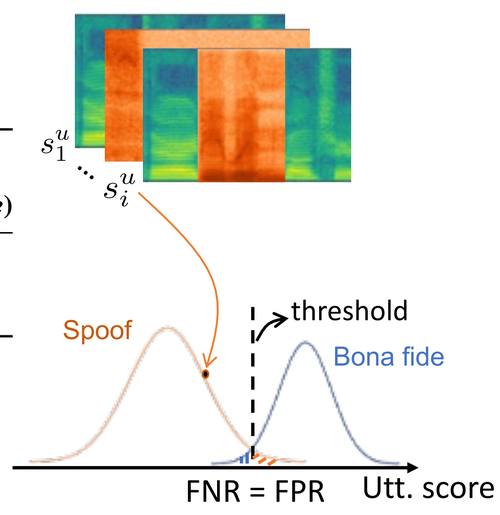
EERs in Partial Spoof

Table 1: *Confusion matrix for spoof localization.*

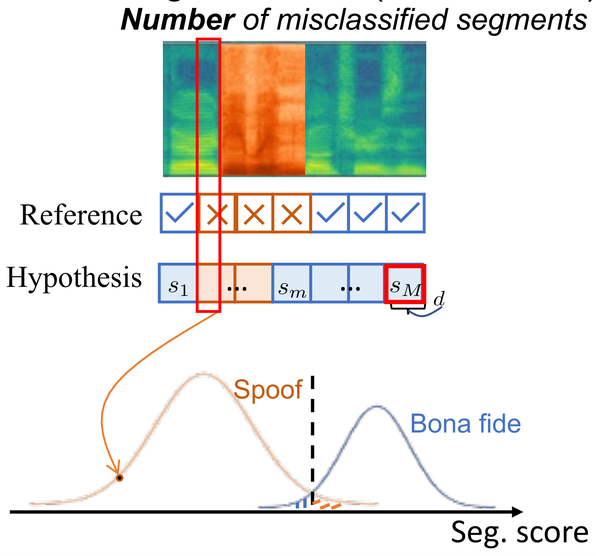
		Hypothesis: Positive (*spoof*)	Hypothesis: Negative (*bona fide*)
Reference	**Positive** ($\mathcal{P}$)	TP	FN
	Negative ($\mathcal{N}$)	FP	TN

Spoof: ☐ ☒ Bona fide: ☐ ☑

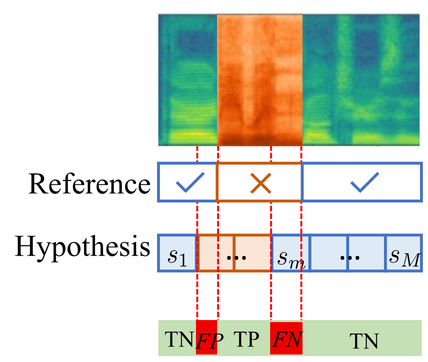
- Utterance EER



- Segment EER (Point-based)
  ***Number** of misclassified segments*



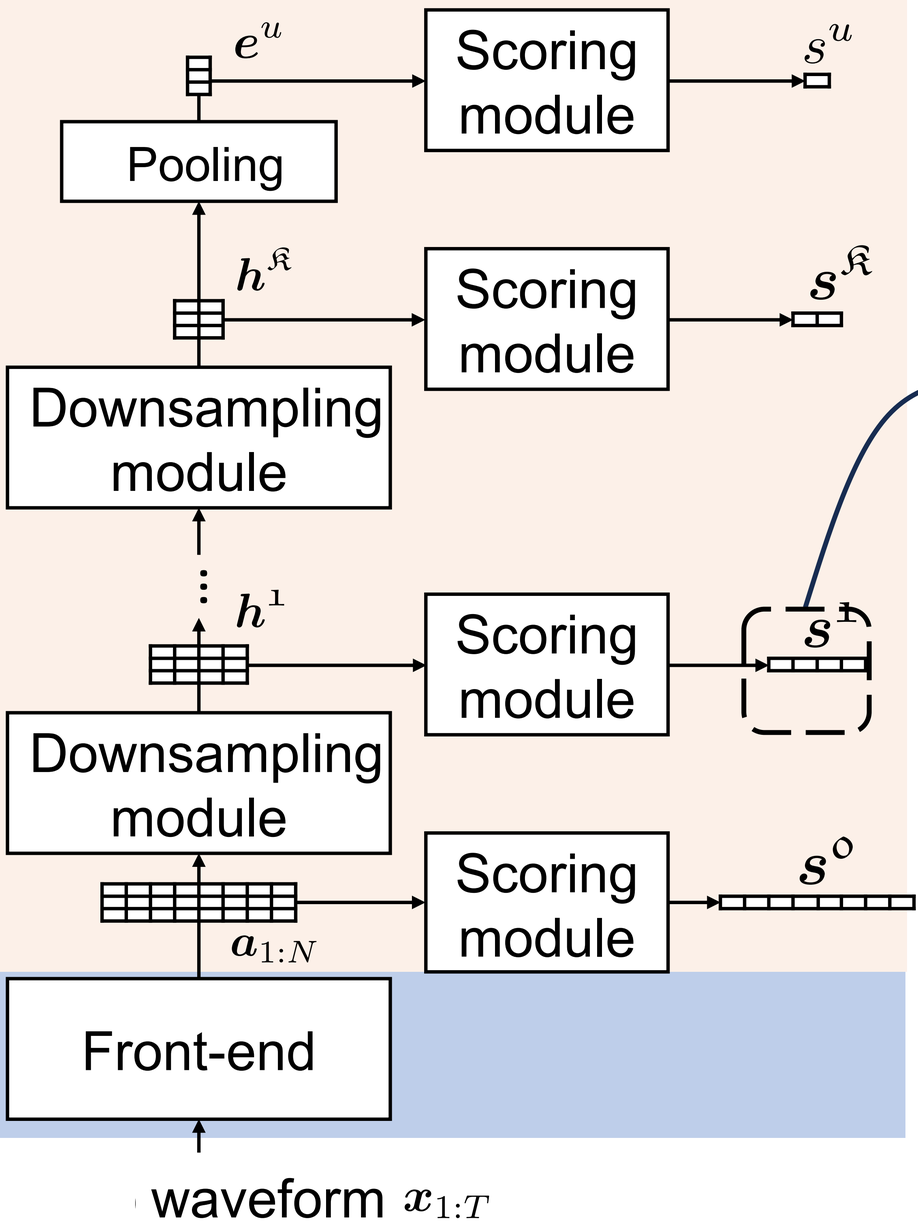
- Range-based EER
  ***Duration** of misclassified regions*



(Please refer to the paper for calculating Range-Based EER with an upgraded binary search algorithm.)

Range-Based EER vs. Point-Based EER

Database: PartialSpoof [1]



Range-based and point-based EER (%) of multi-reso. CM in the eval. set of PartialSpoof.

*Each **row** represents the temporal resolution for **training** and shares the same predicted scores, while each **column** presents the temporal resolution for error **measurement**.)*

Predicted Score	Reso. of Training	Range-based EER	Point-based EER 10	20	40	80	160	320	640
s^0	20	30.40	29.78	*12.84*	11.94	10.52	8.42	5.96	4.06
s^1	40	30.11	29.93	29.94	*11.94*	10.51	8.43	5.98	4.10
s^2	80	30.65	30.12	30.12	30.15	*10.92*	8.70	6.14	4.15
s^3	160	31.36	30.49	30.50	30.52	30.56	*9.24*	6.40	4.11
s^4	320	33.91	33.39	33.38	33.41	33.45	33.48	*6.34*	3.97
s^5	640	37.38	37.53	37.53	37.54	37.56	37.56	37.54	*5.19*

Point-based EER is sensitive to the measurement resolutions.

- **Upper triangular** ◹ : measurement reso. > training reso
- **Diagonal:** measurement reso. = training reso
  - "Underestimation": Smaller value only means the task is easier and does NOT mean that spoof localization is more accurate.
- **Lower triangular** ◺ : measurement reso. < training reso

Range-based EER can measure errors at a finer level.

Conclusion

Using (purple box) **range-based EER**, or point-based EER with unseen and finer temporal resolutions compared with the training resolution. And considering range-based EER when the training temporal resolution is unknown!

References

[1] L. Zhang, X. Wang, E. Cooper, J. Yamagishi, J. Patino, N. Evans. (2021) An Initial Investigation for Detecting Partially Spoofed Audio. Proc. Interspeech 2021, 4264-4268.
[2] L. Zhang, X. Wang, E. Cooper, J. Yamagishi (2021) Multi-task Learning in Utterance-level and Segmental-level Spoof Detection. Proc. ASVspoof workshop 2021, 9-15.
[3] L. Zhang, X. Wang, E. Cooper, N. Evans, J. Yamagishi, "The PartialSpoof Database and Countermeasures for the Detection of Short Fake Speech Segments Embedded in an Utterance," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 31, pp. 813-825, 2023, doi: 10.1109/TASLP.2022.3233236.
[4] This one: Lin Zhang, Xin Wang, Erica Cooper, Nicolas Evans and Junichi Yamagishi, "Range-Based Equal Error Rate for Spoof Localization." (INTERSPEECH 2023)

More about Partial Spoof