



Does Speaker Anonymization Really Work?

CNT 5410: Computer and Network Security

Nigama Annapurna Dendukuri

Kartheek Reddy Gade

Pavan Siva Sai Savaram

Sai Ram Varma Budharaju

Venkata Sai Karthik Metlapalli

Overview

- Introduction
- Background
- Approach
- Analysis of Papers
 - Anonymization models overview
 - Threat model
 - Comparision and selection
- Results
- Challenges
- Conclusion
- Future work

Introduction



- **Speaker anonymization** involves concealing the identity of the speaker, without affecting the intelligibility of the content.
 - Why: Ensures confidentiality by protecting privacy of individuals .
 - How: Can be done using physical or logical techniques.
- A good anonymization technique:
 - Has low similarity to original voice
 - Is intelligible
 - Sounds relatively natural
 - Is difficult to reverse engineer
- But the current voice anonymization methods lack robust security and requires balancing the need for efficient anonymization with processing speed

Background

- Voice Anonymization is the process of obscuring a speaker's identity while preserving speech content.
- Existing anonymization techniques use:
 - Physical techniques:
 - Noise addition
 - Logical techniques:
 - Voice Transformation -> Changing the pitch
 - Voice Conversion -> Converting from male voice to female.
 - Voice Synthesis -> Text to Speech
 - Voice Signal Processing -> Manipulating the acoustic signals



Understanding Related concepts



Key concepts to understand:

Automatic Speaker Verification (ASV): Verify the speaker's identity

Automatic Speech Recognition (ASR): Transcribing speech without identifying the speaker.



Metrics used:

EER (Equal Error Rate): Crucial for ASV, measuring the point where false acceptance equals false rejection.

WER (Word Error Rate): Important for ASR, assessing transcription accuracy.



Voice privacy challenge 2020:

Focused on advancing speaker anonymization systems and evaluating their efficacy.

Involved open-source models and comprehensive criteria, diverse datasets.

Approach



Thoroughly study speaker anonymization methods.



Analyze strengths and vulnerabilities.

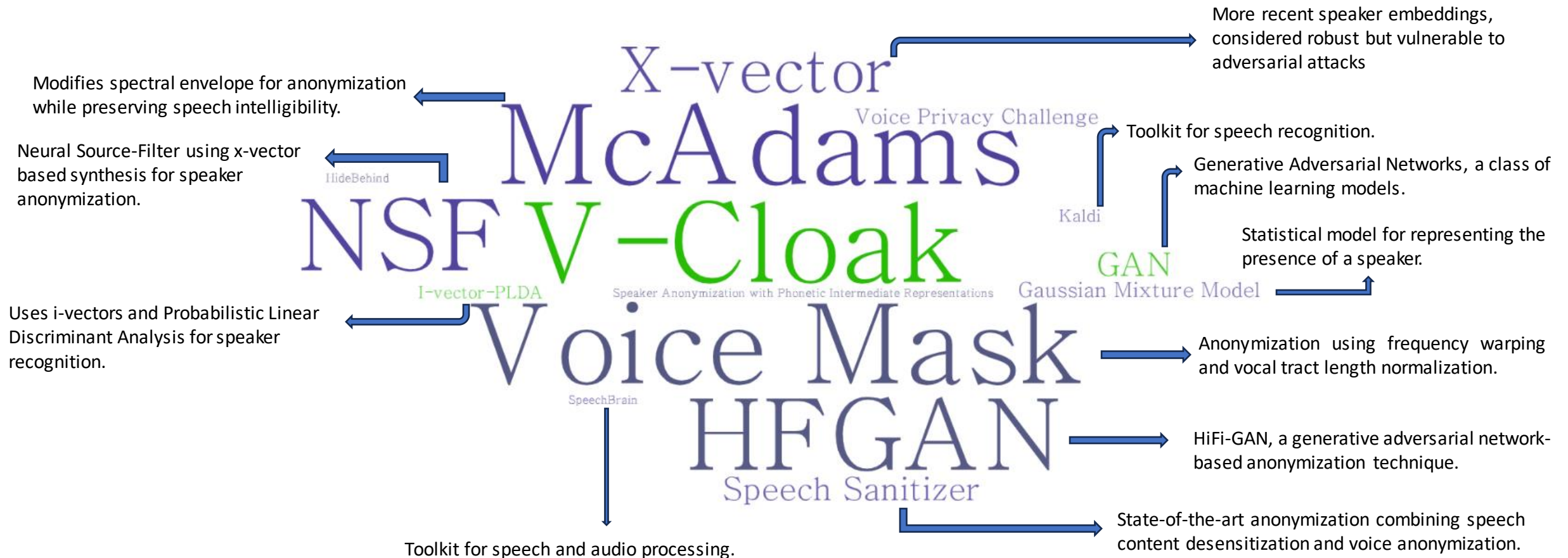


Explore ethical implications.



Engage in discussions about potential misuse and ethical responsibilities.

Related Work



Threat model

- The threat models defined to assess the effectiveness of anonymization techniques are:
 - Ignorant Adversary (A1)
 - Semi-Informed Adversary (A2)
 - Informed Adversary (A3)
- **NSF**: Effective against **A1** adversaries, but it is less effective against A2 and A3 adversaries.
- **HFGAN**: More effective against **A2** adversaries than NSF, but it is still less effective against A3 adversaries.
- **McAdams**: Effective against **A1 and A2** adversaries, but it is less effective against A3 adversaries.
- **VoiceMask**: More effective against **A3** adversaries than NSF, HFGAN, and McAdams.
- **V-Cloak**: Most effective technique against **all three** types of adversaries

Model Selection

- Based on the performance of the models we have decided to perform in-depth analysis on V-cloak
- Real world effectiveness was assessed in ensuring high-level security
- V-cloak was evaluated to determine if it is the latest and most advanced approach.
- The challenges associated with V-cloak was also assessed

Model		B0 (%)	NSF (%)			HFGAN (%)			McAdams (%)			VoiceMask (%)			V-CLOAK (%)		
		EER	MMR	WMR	EER	MMR	WMR	EER	MMR	WMR	EER	MMR	WMR	EER	MMR	WMR	EER
ASV	EP	3.72	88.89	3.89	38.09	87.33	3.89	42.21	46.53	3.89	20.69	70.15	3.89	23.40	97.90	3.89	42.21
	XV	5.74	87.33	4.73	34.47	88.89	4.73	39.05	84.28	4.73	40.19	95.73	4.73	37.79	100.0	4.73	44.73
	DP	3.72	93.97	4.05	39.70	89.39	4.05	33.13	80.15	4.05	35.00	99.24	4.05	41.37	99.77	4.05	49.47
AVG		4.39	90.06	4.22	37.42	88.54	4.22	38.13	70.32	4.22	31.96	88.37	4.22	34.19	99.22	4.22	45.47
WCS		-	87.33	3.89	34.47	87.33	3.89	33.13	46.53	3.89	20.69	70.15	3.89	23.40	97.90	3.89	42.21

AVG: average, **WCS:** worst-case scenario. **EP:** ECAPA-TDNN, **XV:** X-vector, **DP:** DeepSpeaker.

Performance results for the models

Kaldi configure setup success

```
Applications: 3 Thunar 2 Xfce terminal
Terminal - sbudharaju@c38a-s9:/blue/vbindschaedler/sbudharaju/speaker-anonymization/Voice-Privacy-Challenge-2020/kaldi/src
File Edit View Terminal Tabs Help
ps/compilers/cuda/10.0.130
Configuring KALDI to use MKL.
Backing up kaldi.mk to kaldi.mk.bak ...
Checking compiler /home/sbudharaju/.conda/envs/speechbrain/bin/x86_64-conda-linux-gnu-c++ ...
Checking OpenFst library in /blue/vbindschaedler/sbudharaju/speaker-anonymization/Voice-Privacy-Challenge-2020/kaldi/tools/openfst-1.6.7 ...
Checking cub library in /blue/vbindschaedler/sbudharaju/speaker-anonymization/Voice-Privacy-Challenge-2020/kaldi/tools/cub-1.8.0 ...
Doing OS specific configurations ...
On Linux: Checking for linear algebra header files ...
MKL configured with threading: sequential, libs: -L/apps/compilers/intel/2023.2.0.49397/mkl/2023.2.0/lib/intel64 -Wl,-rpath=/apps/compilers/intel/2023.2.0.49397/mkl/2023.2.0/lib/intel64 -lmkl_intel_lp64 -lmkl_core -lmkl sequential
MKL include directory configured as: /apps/compilers/intel/2023.2.0.49397/mkl/2023.2.0/lib/intel64/../../../../include
Configuring MKL threading as
./configure: line 366: libs: bad array subscript
MKL threading libraries configured as -ldl -lpthread -lm
Using Intel MKL as the linear algebra library.
Intel(R) oneAPI Math Kernel Library Version 2023.2-Product Build 20230613 for Intel(R) 64 architecture applications
Successfully configured for Linux with MKL libs from
Using CUDA toolkit /apps/compilers/cuda/10.0.130 (nvcc compiler and runtime libraries)
INFO: Configuring Kaldi not to link with Speex. Don't worry, it's only needed if
you intend to use 'compress-uncompress-speex', which is very unlikely.
../kaldi.mk:49: *** MKLR00T not defined.. Stop.
./configure: line 234: ./exp-test: No such file or directory
Kaldi has been successfully configured. To compile:

make -j clean depend; make -j <NCPU>

where <NCPU> is the number of parallel builds you can afford to do. If unsure,
use the smaller of the number of CPUs or the amount of RAM in GB divided by 2,
to stay within safe limits. 'make -j' without the numeric value may not limit
the number of parallel jobs at all, and overwhelm even a powerful workstation,
since Kaldi build is highly parallelized.
(speechbrain) [sbudharaju@c38a-s9 src]$ make -j clean depend
kaldi.mk:49: *** MKLR00T not defined.. Stop.
(speechbrain) [sbudharaju@c38a-s9 src]$
```

X-vector extraction step in Voice Privacy Challenge

```
Applications | blue | Xfce Terminal | Sat 2 Dec, 19:15 Sai Budharaju
Terminal - sbudharaju@c38a-s9:/blue/vbindschaedler/sbudharaju/speaker-anonymization/Voice-Privacy-Challenge-2020/baseline
File Edit View Terminal Tabs Help

Done

Stage 5: Downloading LibriTTS data sets for training anonymization system (train-other-500)...
local/download_and_untar.sh: data part train-other-500 was already successfully extracted, nothing to do.

Stage 6: Prepare anonymization pool data...
utils/fix_data_dir.sh: file data/libritts_train_other_500/utt2spk is not in sorted order or not unique, sorting it
utils/fix_data_dir.sh: file data/libritts_train_other_500/spk2utt is not in sorted order or not unique, sorting it
utils/fix_data_dir.sh: file data/libritts_train_other_500/text is not in sorted order or not unique, sorting it
utils/fix_data_dir.sh: file data/libritts_train_other_500/wav.scp is not in sorted order or not unique, sorting it
utils/fix_data_dir.sh: file data/libritts_train_other_500/spk2gender is not in sorted order or not unique, sorting it
fix_data_dir.sh: kept all 205044 utterances.
fix_data_dir.sh: old files are kept in data/libritts_train_other_500/.backup
utils/validate_data_dir.sh: Successfully validated data-directory data/libritts_train_other_500
local/data_prep_libritts.sh: successfully prepared data in data/libritts_train_other_500

Stage 7: Extracting xvectors for anonymization pool...
steps/make_mfcc.sh --write-utt2num-frames true --mfcc-config conf/mfcc.conf --nj 3 --cmd run.pl data/libritts_train_other_500 exp/make_mfcc /blue/vbindschaedler/sbudharaju/speaker-anonymization/Voice-Privacy-Challenge-2020/baseline/mfcc
utils/validate_data_dir.sh: Successfully validated data-directory data/libritts_train_other_500
steps/make_mfcc.sh: [info]: no segments file exists: assuming wav.scp indexed by utterance.
steps/make_mfcc.sh: Succeeded creating MFCC features for libritts_train_other_500
fix_data_dir.sh: kept all 205044 utterances.
fix_data_dir.sh: old files are kept in data/libritts_train_other_500/.backup
sid/compute_vad_decision.sh --nj 3 --cmd run.pl data/libritts_train_other_500 exp/make_vad /blue/vbindschaedler/sbudharaju/speaker-anonymization/Voice-Privacy-Challenge-2020/baseline/mfcc
Created VAD output for libritts_train_other_500
fix_data_dir.sh: kept all 205044 utterances.
fix_data_dir.sh: old files are kept in data/libritts_train_other_500/.backup
sid/nnet3/xvector/extract_xvectors.sh --cmd run.pl --nj 3 exp/models/2_xvect Extr/exp/xvector_nnet_1a data/libritts_train_other_500 exp/models/2_xvect Extr/exp/xvector_nnet_1a/anon/xvectors libritts_train_other_500
sid/nnet3/xvector/extract_xvectors.sh: using exp/models/2_xvect Extr/exp/xvector_nnet_1a/extract.config to extract xvectors
sid/nnet3/xvector/extract_xvectors.sh: extracting xvectors for data/libritts_train_other_500
sid/nnet3/xvector/extract_xvectors.sh: extracting xvectors from nnet
```


Output of Voice Privacy challenge ASR - McAdams

```
Terminal - sbudharaju@c38a-s5:/blue/vbindschaedler/sbudharaju/speaker-anonymization/Voice-Privacy-Challenge-2020/baseline
File Edit View Terminal Tabs Help
5772 / 5774
5773 / 5774
5774 / 5774

Stage 10: Making VCTK anonymized evaluation subsets...
utils/subset_data_dir.sh: reducing #utt from 5766 to 5422
utils/subset_data_dir.sh: reducing #utt from 5766 to 344
utils/subset_data_dir.sh: reducing #utt from 5606 to 5255
utils/subset_data_dir.sh: reducing #utt from 5606 to 351
utils/subset_data_dir.sh: reducing #utt from 5674 to 5328
utils/subset_data_dir.sh: reducing #utt from 5674 to 346
utils/subset_data_dir.sh: reducing #utt from 5774 to 5420
utils/subset_data_dir.sh: reducing #utt from 5774 to 354

Stage 11: Evaluate datasets using speaker verification...
**ASV: libri_dev_trials_f, enroll - original, trial - original**
ASV scoring: exp/results-2023-12-05-00-46-00/ASV-libri_dev_enrolls-libri_dev_trials_f
EER: 8.665%
minDCF(p-target=0.01): 0.4826
minDCF(p-target=0.001): 0.5412
Cllr (min/act): 0.304/42.926
ROCCH-EER: 8.571%

linkability: 0.799043

libri_dev_enrolls-libri_dev_trials_f
Population: 0.492 bit
Individual: 3.826 (C)

**ASV: libri_dev_trials_f, enroll - original, trial - anonymized**
compute MFCC: libri_dev_trials_f_anon
steps/make_mfcc.sh --nj 3 --cmd run.pl --write-utt2num-frames true data/libri_dev_trials_f_anon
steps/make_mfcc.sh: moving data/libri_dev_trials_f_anon/feats.scp to data/libri_dev_trials_f_anon/.backup
utils/validate_data_dir.sh: Successfully validated data-directory data/libri_dev_trials_f_anon
```

Results

- **Equal Error Rate (EER)**
 - 8.65%
- **Minimum Detection Cost Function (minDCF)**
 - 0.4826 for $p(\text{target}=0.01)$
 - 0.5412 for $p(\text{target}=0.001)$
- **Clustering Identification Rate (CIIR)**
 - 0.3044 (min action)
 - 0.4296 (overall)
- **Receiver Operating Characteristic Curve - Equal Error Rate (ROCCH-EER)**
 - 8.57%
- **Linkability**
 - 0.799043

Challenges



HARDWARE LIMITATIONS: REQUIRED GPU
IN MANY MODELS. NEEDED ACCESS TO
HIPERGATOR



**TOOLKIT AND SOFTWARE
DEPENDENCIES:** KALDI AND SPEECHBRAIN
WERE COMPLEX AND RESOURCE INTENSIVE



LINUX COMPATIBILITY: REQUIRED DUAL
BOOTING THE SYSTEM.

Conclusion

- **Speaker Verification Effectiveness**
 - EER and ROCCH-EER rates indicate moderate system performance in verification tasks.
- **Accuracy and Cost Trade-off**
 - minDCF values point to a reasonable misclassification cost, highlighting potential for optimization.
- **Linkability vs Anonymization**
 - High linkability score suggests effective speaker linking, yet poses a challenge for stronger anonymization.
- **Key Takeaway**
 - The system shows promise but requires advancements in balancing verification accuracy with anonymization needs.

Future Work

Deep Learning Models

- Utilize advanced models trained on diverse datasets for improved accuracy.

• Adversarial Training

- Train systems against adversarial attacks for enhanced security.

• User-Centric Design

- Focus on user-friendly interfaces and feedback for practical use.

• Ethical and Legal Considerations

- Address ethical implications and adhere to privacy regulations.

• Collaboration with Research Community

- Engage with challenges and incorporate community-driven advancements.

Thank You!