Identifying Vulnerabilities in Voice Biometric Systems

By Jay Shah (1461-3930), Shalaka Vinod Naik (9139-1915), Hareesh Nutalapati (6801-1198)

MOTIVATION

Speaker Verification have been substantially advanced in the past decades. Despite advances, these systems still lack robustness: their performance degrades dramatically when the acoustic training data is mismatched to the test conditions. In this study we will perform speech synthesis attacks on such biometric system to identify its vulnerability. We will formalize the space of adversary and introduce techniques which will force misclassification by traditional voice recognition system. Our results shows that these system are largely ineffective to our attacks. Our aim is to expose the weak point of a system so that we can further delve into a cause and rectify the problems.

RELATED WORK

Speaker recognition refers to recognizing persons from their voice. No two individuals sound identical because of the difference in their voice production organs and characteristic manner of speaking. Speaker recognition systems use several these features in parallel, attempting to cover these different aspects and employing them in a complementary way to achieve more accurate recognition. The existing voice biometric system can be made accurate by implementing the technique of score. This method introduces a new strategy called test normalization based on the mean and variance estimation. In early studies speaker models were generated by time-averaging features so that each utterance could be represented as a single vector. The average vectors would then be compared using a distance measure which is computationally very efficient but gives poor recognition accuracy. A recently discovered robust method is to present utterances using a single vector, a so-called supervector. Although these techniques have been implemented, several attacks have managed to break the voice biometric system. Our aim is to carry out such attacks on a speaker verification system and compute the results accordingly.

TOOLS

- **Speaker Verification System**
  Open source Bob Spear Speaker Verification System
- .**Voice Conversion System**
  Festvox voice coversion system.

OVERVIEW

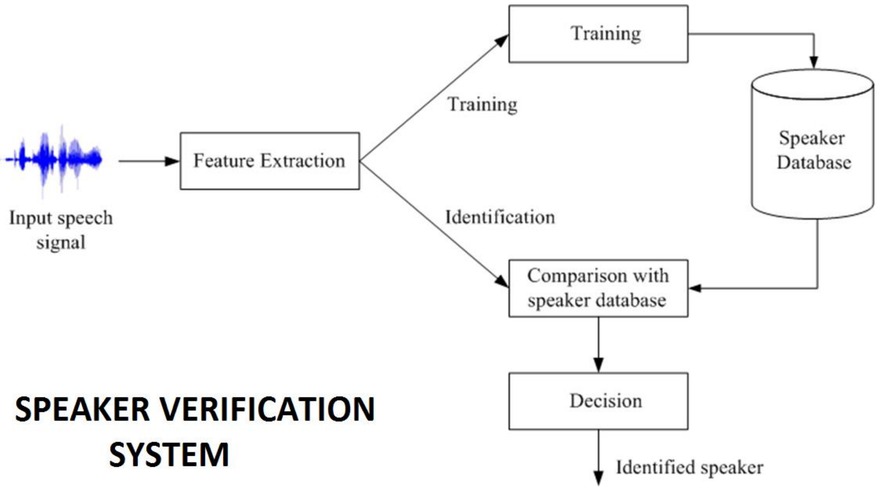
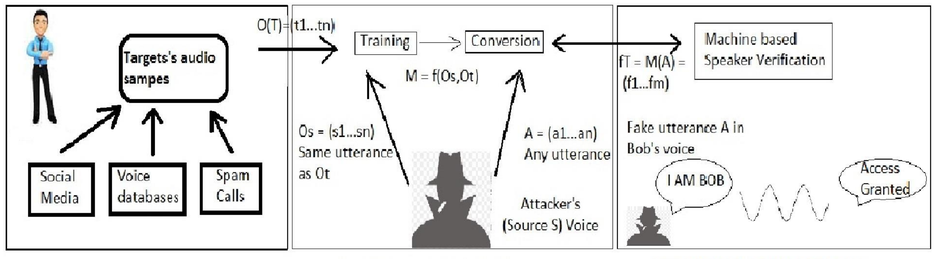


Figure 1

EXPERIMENTAL APPROACH



ATTACK ON SPEAKER VERIFICATION SYSTEM

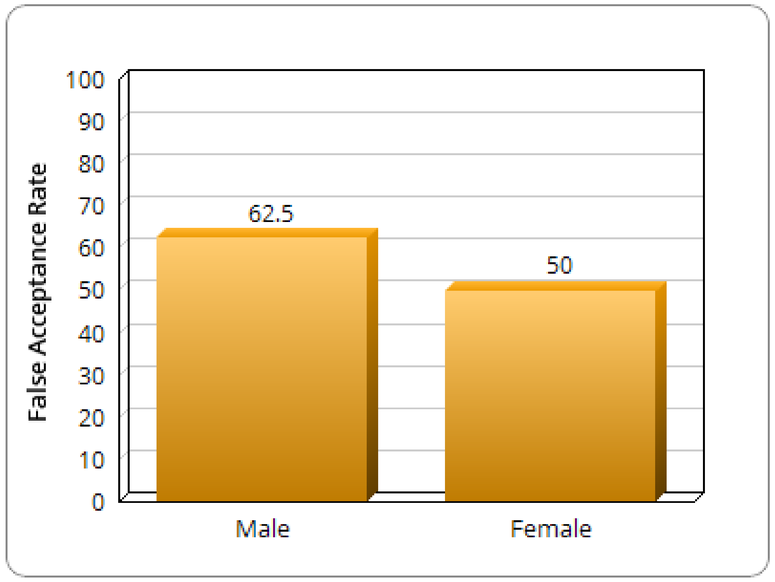
Figure 2

TECHNIQUES AND METRICS USED

- UBM-GMM: It is a modelling technique that uses spectral features and computes log-likelihood of the Gaussian Mixture Models for speaker verification system.
- The number of false positives by the generated by the system in response to the converted voice samples were recorded.
- Metrics Used: The performance of a speaker verification system is evaluated based on the False Acceptance Rates (FAR).

RESULTS

Modelling Technique	FAR	
	Male	**Female**
UBM-GMM	62.5%	50.0%



CONCLUSIONS

- Human voices can be easily stolen and used to access confidential info of an entity. We showed that voice conversion poses a serious threat and our attacks can be successful for a majority of cases. New voice morphing techniques are used and these are getting better by the day.
- The best way to prevent such an attack in the first place is by not allowing the imposter to get access to your voice samples. This can be done by spreading awareness.