

**Doctoral School of Information and Biomedical Technologies
Polish Academy of Sciences (TIB PAN)**

SUBJECT:

Analysis of speech credibility and voice signal reliability for detecting suspicious user intentions

SUPERVISOR:

dr hab. inż. Grzegorz Borowik
grzegorz.borowik@nask.pl
NASK-PIB, Centrum Badań i Rozwoju
Kolska 12, 01-045 Warszawa

DESCRIPTION:

The doctoral thesis will focus on analyzing the speech credibility and reliability of voice signals for the purpose of detecting suspicious user intentions. It will involve identifying voice features and patterns that indicate potentially fraudulent behavior. The research will employ machine learning models that can detect fraud based on audio data while being resilient to attacks.

The research will investigate the use of voice analysis to improve fraud detection techniques, which are increasingly important in many fields, including finance, security, and law enforcement. The proposed methods will aim to enhance the accuracy and effectiveness of fraud detection systems by analyzing voice data, which provides a rich source of information about the user. The models developed in this study will help reduce the incidence of fraud and improve the overall security of systems that rely on voice authentication.

Overall, this doctoral thesis will provide a valuable contribution to the field of voice analysis and fraud detection. By developing resilient machine learning models, the research will show how voice data can be used to improve the accuracy of fraud detection systems. The findings of this study will have practical applications in many domains and will help prevent fraudulent activities that pose a threat to individuals and organizations.

BIBLIOGRAPHY:

1. Lavrentyeva, G., Volkova, M., Avdeeva, A., Novoselov, S., Gorlanov, A., Andzhukaev, T., Ivanov, A., & Kozlov, A. (2020). Blind Speech Signal Quality Estimation for Speaker Verification Systems. *Inter-speech*.
2. Rodríguez-Ortega, Y., DoraMaríaBallesteros, L., & Renza, D. (2020). A Machine Learning Model to Detect Fake Voice. *International Conference on Applied Informatics*.
3. Gao, Y., Zheng, W., Yang, Z., Kohler, T., Fuegen, C., & He, Q. (2020). Interactive text-to-speech via semi-supervised style transfer learning.

4. Jalal, M., Mays, K. K., Guo, L., & Betke, M. (2020). Performance Comparison of Crowdworkers and NLP Tools on Named-Entity Recognition and Sentiment Analysis of Political Tweets.
5. Li, J., Sun, A., Han, J., & Li, C. (2020). A Survey on Deep Learning for Named Entity Recognition. *IEEE transactions on knowledge and data engineering*.
6. Chettri, B., Kinnunen, T., & Benetos, E. (2020). Deep Generative Variational Autoencoding for Replay Spoof Detection in Automatic Speaker Verification.
7. Park, S., Kim, D., & Joe, M. (2020). Cotatron: Transcription-Guided Speech Encoder for Any-to-Many Voice Conversion without Parallel Data.
8. Tanaka, M., Nose, T., Kanagaki, A., Shimizu, R., & Ito, A. (2020). Scyclone: High-Quality and Parallel-Data-Free Voice Conversion Using Spectrogram and Cycle-Consistent Adversarial Networks.
9. Guo, H., Zhang, S., Soong, F. K., He, L., & Xie, L. (2020). Conversational End-to-End TTS for Voice Agent.
10. Jia, Y., & SungChu, S. (2020). A Deep Learning System for Sentiment Analysis of Service Calls.
11. Łańcucki, A., et al. (2020). Robust Training of Vector Quantized Bottleneck Models.
12. Reimao, R. (2019). Synthetic Speech Detection Using Deep Neural Networks. Master's Thesis, Graduate Program in Computer Science and Engineering, York University.
13. Raghavan, P., & Gayar, N.E. (2019). Fraud Detection using Machine Learning and Deep Learning. 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), 334-339.
14. Wang, S., Yuan, W., Wang, J., & Unoki, M. (2019). Detection of speech tampering using sparse representations and spectral manipulations based information hiding. *Speech Commun.*, 112, 1-14.
15. Tripathi, S., Kumar, A., Ramesh, A., Singh, C., & Yenigalla, P. (2019). Deep Learning based Emotion Recognition System Using Speech Features and Transcriptions.
16. Lian, Z., Tao, J., Liu, B., & Huang, J. (2019). Unsupervised Representation Learning with Future Observation Prediction for Speech Emotion Recognition.
17. Gao, Y., Singh, R., & Raj, B. (2018). Voice Impersonation Using Generative Adversarial Networks. Electrical and Computer Engineering Department, Carnegie Mellon University.
18. Niu, Y., Zou, D., Niu, Y., He, Z., & Tan, H. (2017). A breakthrough in Speech emotion recognition using Deep Retinal Convolution Neural Networks.
19. Safavi, S., Gan, H.C., Mporas, I., & Sotudeh, R. (2016). Fraud Detection in Voice-Based Identity Authentication Applications and Services. 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), 1074-1081.
20. Faúndez-Zanuy, M., Hagsmüller, M., & Kubin, G. (2007). Speaker identification security improvement by means of speech watermarking. *Pattern Recognit.*, 40, 3027-3034.
21. Mporas, I., Ganchev, T., Siafarikas, M., & Fakotakis, N. (2007). Comparison of Speech Features on the Speech Recognition Task. *Journal of Computer Science*, 3(8), 608-616.
22. Grosser, H., Britos, P., & García-Martínez, R. (2005). Detecting Fraud in Mobile Telephony Using Neural Networks. *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*.