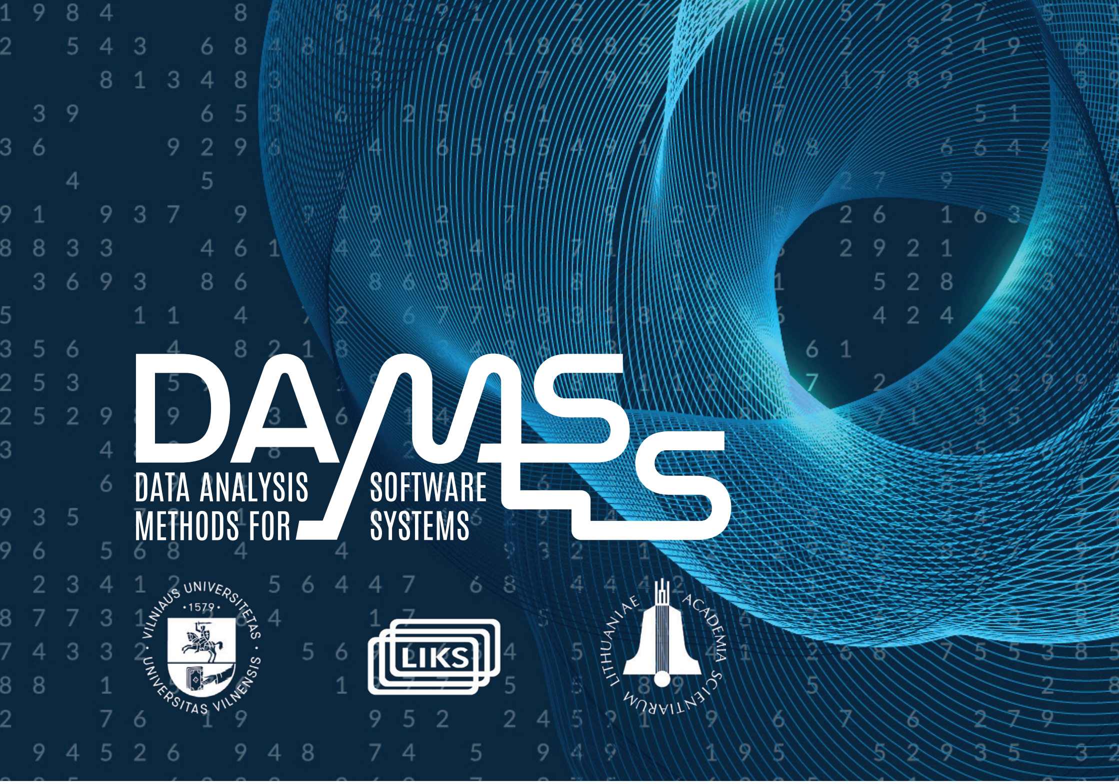

15th Conference on

DATA ANALYSIS METHODS for Software Systems

November 28–30, 2024

Druskininkai, Lithuania, Hotel "Europa Royale"
https://www.mii.lt/DAMSS

Co-Chairmen:

Prof. **Gintautas Dzemyda** (Vilnius University, Lithuanian Academy of Sciences)
Dr. **Saulius Maskeliūnas** (Lithuanian Computer Society)

Programme Committee:

Dr. **Jolita Bernatavičienė** (Lithuania)
Prof. **Juris Borzovs** (Latvia)
Prof. **Janis Grundspenkis** (Latvia)
Prof. **Janusz Kacprzyk** (Poland)
Prof. **Ignacy Kaliszewski** (Poland)
Prof. **Bożena Kostek** (Poland)
Prof. **Tomas Krilavičius** (Lithuania)
Prof. **Olga Kurasova** (Lithuania)
Assoc. Prof. **Tatiana Tchemisova** (Portugal)
Assoc. Prof. **Gintautas Tamulevičius** (Lithuania)
Prof. **Julius Žilinskas** (Lithuania)

Organizing Committee:

Dr. **Jolita Bernatavičienė**
Prof. **Olga Kurasova**
Assoc. Prof. **Viktor Medvedev**
Laima Paliulionienė
Assoc. Prof. **Martynas Sabaliauskas**
Prof. **Povilas Treigys**

Contacts:

Dr. Jolita Bernatavičienė
jolita.bernataviciene@mif.vu.lt
Tel. (+370 5) 2109 315
Prof. Olga Kurasova
olga.kurasova@mif.vu.lt

https://doi.org/10.15388/DAMSS.15.2024
ISBN 978-609-07-1112-5 (digital PDF)

Are We Able to Model the Human Auditory System in Speech Signal Enhancement?

Daniel Zakševski, Gintautas Tamulevičius

Institute of Data Science and Digital Technologies
Vilnius University

gintautas.tamulevicius@mif.vu.lt

Human hearing has unique characteristics that are actively modeled and integrated into speech signal analysis and processing. Masking, frequency selectivity, individual nature, and adaptivity of the human auditory system are the features we are trying to understand, model, and implement in speech signal processing applications. In the speech enhancement domain, the human auditory system is often modeled by band-pass filters, and now artificial neural networks are being used increasingly. While many new ideas and models are being proposed and developed, most have inherent weaknesses. Some solutions are based on predefined and static filters (e. g., mel frequency cepstral analysis, bark scale filters, and equivalent rectangular bandwidth filters). This is inconsistent with the adaptive and individual nature of the human auditory system, as these static models cannot capture the dynamic changes in auditory filter shapes or the individual differences in hearing abilities. Other solutions are considered data-driven, i.e., a neural network trained on specific data is proposed as a method. A persistent challenge in current speech enhancement methods is their significant performance degradation in low signal-to-noise ratio (SNR) environments. This contrasts with the human auditory system's ability to effectively extract and comprehend speech even under such conditions.

In this study, we are exploring and comparing human auditory models, searching for innovative ideas of dynamic, adaptive, or active properties applicable to speech signal enhancement tasks. Successful modeling of these properties and their integration into non-linear neural models may lead to the development and implementation of a highly efficient, human auditory system-based speech enhancement approach. This will potentially improve the performance of speech enhancement systems in real-world scenarios, leading to more intelligible speech for various applications, such as hearing aids, telecommunications, and voice assistants.