

**JOHAN ANDREAS ROHDIN**

## **SEQUENCE Summarizing neural networks for speaker recognition**

Automatické rozpoznávání mluvčího má za úkol ověřit nebo identifikovat osobu na základě nahrávky. Komerční aplikace rozpoznávání mluvčího jsou v přístupových systémech a v bankovníctví, důležité je také užití je v oblasti národní bezpečnosti.

Navrhovaný projekt se týká automatického rozpoznávání mluvčího a je inspirován obrovským pokrokem, který v několika minulých letech přinesly do detekčních a rozpoznávacích úloh umělé neuronové sítě (neural networks, NN). Cílem našeho projektu je vyvinout nový typ NN, která bude vhodná pro rozpoznávání mluvčího a dosáhnout stavu, kdy bude tento přístup použitelný pro praktické aplikace. Dosavadní snahy použít NN pro rozpoznávání mluvčího se vždy zaměřovaly jen na to, že se jeden nebo několik funkčních bloků v klasických systémech nahradil za NN. Tento přístup je ale omezen (podobně jako existující systémy) tím, že je pevně předepsáno, co má neuronová síť modelovat. Náš projekt se zaměřuje na náhradu celého řetězce bloků v rozpoznávacím systému jednou neuronovou sítí, která bude celé promluvy zpracovávat v jediném kroku. To by mělo lépe využít schopností NN modelovat složité vzory v řečových signálech.

Cíle projektu budou dosaženy teoretickou prací (odvození struktury NN, odvození kritérií pro trénování), implementací (paralelizace, škálování, atd.) a důkladným testováním na reálných řečových datech.

Navrhovaný projekt je plně v souladu s Regionální inovační strategií Jihomoravského kraje, a jeho specializační strategií "Smart Specialization strategy (S3)". S3 definuje výzkum a vývoj v oblasti počítačového hardware a software jako jednu z pěti strategických priorit, a zvláště zmiňuje oblasti IT bezpečnosti a cognitroniky. Navrhovaný projekt zcela spadá do oblasti cognitroniky a vzhledem k hlavním aplikacím rozpoznávání mluvčího se týká i oblasti počítačové bezpečnosti.