

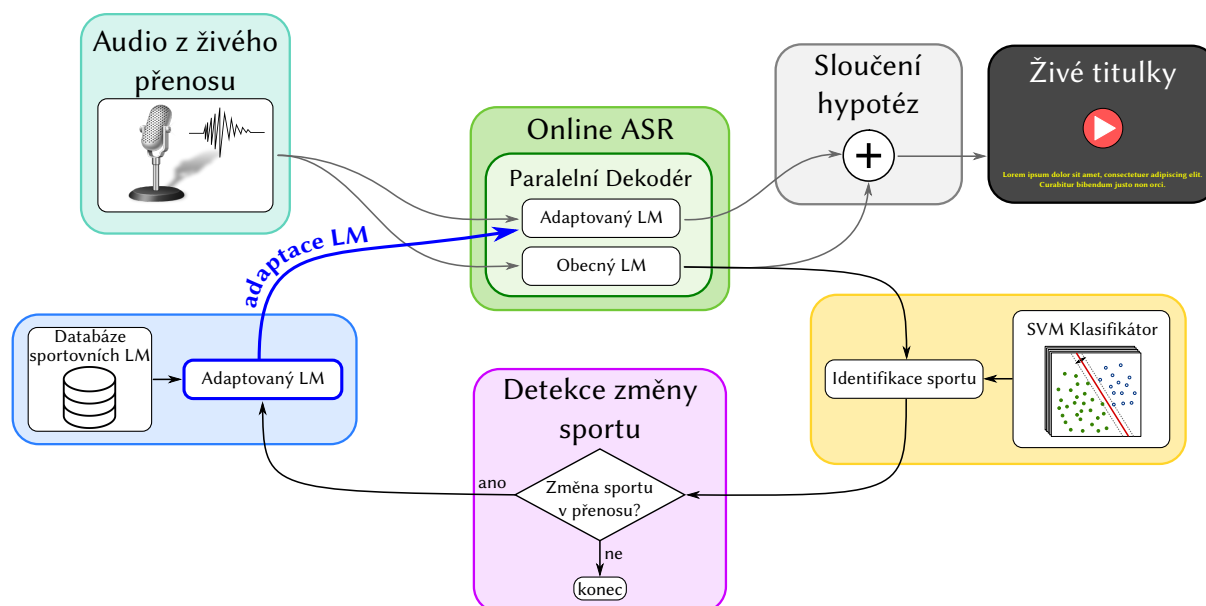
## Adaptace jazykového modelu v úloze automatického titulkování živě vysílaných sportovních přehledů

Jan Lehečka<sup>1</sup>

### 1 Úvod

Při automatickém rozpoznávání řeči (angl. Automatic Speech Recognition, ASR) se využívají tzv. jazykové modely (angl. Language Model, LM), které popisují pravděpodobnosti slov a posloupností slov (tzv. n-gramů) v jazyce. Jazykový model je typicky trénován z velkého množství textů, které jsou blízké doméně (tj. stylu a tématu promluvy), na kterou má být rozpoznávání řeči nasazeno. Každá doména však vyžaduje specifický jazykový model, který co nejpřesněji reflektuje příslušnou slovní zásobu a vazby mezi slovy. V úlohách, kde dochází k časté změně domény, je obvyklé používat kombinaci automatické identifikace domény a adaptace jazykového modelu, která přizpůsobuje jazykový model dle obsahu rozpoznané řeči, viz např. Echeverry-Correa et al. (2015).

V tomto příspěvku je popsáno schéma adaptace jazykového modelu, které je, na rozdíl od běžně používaných přístupů, možné použít pro rozpoznávání řeči v reálném čase (tzv. online). Přínos navrženého systému je demonstrován na úloze živého titulkování sportovních přehledů z olympijských her. Tato úloha je z hlediska ASR velmi obtížná, protože každý sportovní přehled typicky obsahuje značné množství jmen sportovců a sportovkyň (zejména cizinců s komplikovanou výslovností) a během rozpoznávání dochází k častým a rychlým změnám domény.



Obrázek 1: Schéma navrženého systému

<sup>1</sup> student doktorského studijního programu Aplikované vědy a informatika, obor Kybernetika, e-mail: jlehecka@kky.zcu.cz

## 2 Schéma adaptace LM

Schéma celého systému je zobrazeno na obrázku 1. Vstupem je audio signál (živý stream z televizního studia) a výstupem titulky generované v reálném čase. Vstupní signál zpracovává modul pro online rozpoznávání řeči (online ASR), který generuje dvě částečné hypotézy o tom, co bylo zatím řečeno: jednu za použití obecného LM (trénovaného ze všech textů) a druhou s adaptovaným LM (trénovaným cíleně pro konkrétní sport). Hypotézy z obecného LM jsou použity pro identifikaci aktuálního sportu pomocí SVM klasifikátoru. Identifikované sporty jsou monitorovány detektorem změny sportu v promluvě a jakmile je změna detekována, vybere se z databáze vhodný adaptovaný LM, který se za běhu importuje do rozpoznávače. Systém je tak schopný reagovat poměrně rychle na změny domény a za běhu se adaptovat.

## 3 Výsledky

Systém byl otestován na záznamech přehledů České televize z olympijských her 2008 v Pekingu. Celkem bylo rozpoznáno 6,3 hodin záznamů, které byly následně porovnány se správným přepisem pro vyhodnocení chybovosti (angl. Word Error Rate, WER). Protože správné rozpoznání jmen sportovců je klíčové pro dobrou srozumitelnost titulků, byla kromě celkové chybovosti vyhodnocena také chybovost rozpoznávání jmen, tj. kolik procent jmen sportovců bylo rozpoznáno s chybou. Všechna jména sportovců byla přidána do LM ze soupisky pomocí tzv. expanze třídivých jazykových modelů. Výsledky jsou shrnuty v tabulce 1.

	WER (celkem)	WER (jména)
bez adaptace LM	4,94%	24,89%
s adaptací LM	4,42% (-10,5% rel.)	19,73% (-20,7% rel.)

**Tabulka 1:** Chybovost (WER) živých titulků.

## 4 Závěr

Výsledky experimentu potvrzují, že v úloze automatického titulkování živě vysílaných sportovních přehledů přináší online adaptace LM výrazné zlepšení chybovosti (rel. o 10,5%), zejména pak chybovosti rozpoznávání vlastních jmen sportovců (rel. o 20,7%), která jsou klíčová pro porozumění rozpoznávaného textu. Nespornou výhodou navrženého systému je také možnost nasadit jej v budoucnu na živě vysílané přehledy libovolné sportovní akce, je-li předem známa soupiska zúčastněných sportovců.

## Poděkování

Tento příspěvek byl podpořen grantovým projektem SGS-2016-039.

## Literatura

Echeverry-Correa, J.D., Ferreiros-López, J., Coucheiro-Limeres, A., Córdoba, R. a Montero, J.M. (2015) Topic identification techniques applied to dynamic language model adaptation for automatic speech recognition. *Expert Systems with Applications*, Volume 42, pp. 101–112.