

Oponentský posudek diplomové práce

Petra Galuščáková :

Evaluační metody systémů pro vyhledávání v nesegmentované mluvené řeči

Úkolem systémů pro vyhledávání v nesegmentované mluvené řeči je automatická identifikace tématicky koherentních pasáží v záznamech mluvené řeči. Vstupem je specifikace tématu, resp. dotaz, výstupem časově určené začátky a konce relevantních pasáží. Oproti klasickému vyhledávání v kolekcích dokumentů se v tomto případě jedná o vyhledávání na spojitém prostoru (časové ose nahrávek), kde řeč není předem segmentována na kratší úseky („dokumenty“). Řešením úlohy je nalezení relevantního úseku, resp. relevantních úseků, přičemž začátek a konec takového úseku se může nacházet v jakémkoliv okamžiku nahrávky.

Předložená práce studuje vhodný přístup pro vyhodnocování vyhledávání v nesegmentované mluvené řeči. V současné době pro evaluaci výsledků vyhledávání v nesegmentované řeči existuje pouze jediná metrika – a to metrika mGAP (mean Generalized Average Precision). Tato práce se snaží ověřit vhodnost této metriky a navrhnout její případné vylepšení směrem ke zlepšení kvality samotného vyhledávání.

Způsob vyhodnocování, který je v této práci navrhnout, je také empiricky ověřen na trénovacích datech s reálnými uživateli. Je sledováno chování uživatelů/respondentů při vyhledávání, a to při různých situacích, které při automatickém vyhledávání nastávají. Jako testovací kolekce je použita kolekce nahrávek výpovědí svědků holocaustu, která byla pořízena v rámci mezinárodního projektu Malach (Multilingual Access to Large Spoken Archives) v letech 1994 – 1997 nadací Shoah Visual History Foundation.

Úvodní část práce je věnována jemnému úvodu do oblasti information retrieval, a konkrétněji též oblasti vyhledávání informací v mluvené řeči. Jsou zde také stručně charakterizovány jednotlivé techniky a systémy, které jsou v současné době využívány pro oblast information retrieval. Druhá kapitola je věnována metodám a metrikám pro vyhodnocování kvality vyhledávání, tyto metriky jsou zde celkem podrobně popsány, včetně jejich výhod a nevýhod. Důraz je při analýze kladen především na metriku mGAP (mean Generalized Average Precision), která byla navržena zejména pro oblast vyhledávání v nesegmentované řeči. Kapitola třetí, která je jádrem celé práce, se věnuje návrhu na vytvoření nového způsobu vyhodnocování vyhledávání v nesegmentované mluvené řeči. Jednou z úloh této práce je ověřit, zda použití metriky mGAP je opodstatněné, a v případě, že by byly nalezeny nedostatky této metriky, navrhnout adekvátní řešení. V rámci ověření adekvátnosti metriky mGAP byl uskutečněn uživatelský průzkum. V rámci výzkumu bylo na simulovaném vyhledávacím systému vyhodnocováno uživatelské chování respondentů – sledována byla zejména závislost času a spokojenosti uživatelů se vzdáleností a umístěním bodu vyhledaného vyhledávacím systémem a skutečným začátkem tématu v nahrávce. Pro účely průzkumu autorka vytvořila vlastní přehrávač dat, který respondentům přehrává jednotlivé nahrávky tak, aby respondent mohl co možná nejrychleji nalézt začátek relevantního úseku v nahrávce ke konkrétnímu tématu, které je respondentovi předloženo. V dalších dvou kapitolách jsou popsány výsledky průzkumu a jeho průběh. Je zde celkem podrobně popsáno uživatelské rozhraní programu/přehrávače – jeho funkcionality, ovládání, ale i implementace. Kapitola šestá obsahuje konečný návrh metriky a její porovnání s existujícími metrikami pro evaluaci vyhledávání v mluvené řeči. Na základě průzkumu byla jen mírně pozměněna penalizační funkce metriky mGAP. Nicméně výsledky výzkumu

potvrdily použitelnost původní mGAP metriky pro evaluaci vyhledávání v nesegmentované mluvené řeči.

Cílem práce bylo ověřit, zda jsou v současnosti používané způsoby vyhledávání v nesegmentované mluvené řeči adekvátní, případně navrhnout možnosti zlepšení hodnocení těchto systémů. V práci jsou využity zejména empirické přístupy založené na tom, jak uživatelé vyhledávacích systémů pro vyhledávání v nesegmentované mluvené řeči vnímají a zejména to, jak s takovými systémy pracují. A v neposlední řadě také to, jaké je subjektivní uživatelské hodnocení výsledků vyhledávání. Na základě získaných výsledků je pak navržena možnost vylepšení používané metriky (mGAP). Upravený způsob evaluace je pak porovnán s původními technikami.

Všechny stanovené cíle autorka ve své práci splnila.

Autorka prokázala schopnost formulovat myšlenky, navrhnout řešení zadaného problému, kriticky jej zhodnotit a navrhnout různá vylepšení. Práce celkově působí uceleně. Členění práce je logické. Množství prostudované literatury je dostatečné. Jazykové zpracování dobré, bez překlepů .

Nedostatky:

s. 20 a 22 - drobná nepřesnost v terminologii – narrative a nikoliv o narrator.

s. 16 – chybně uvedená čísllice. Českých nahrávek bylo cca 360 a nikoliv 700.

Hodnocenou práci považuji za zdařilou. Hodnocená práce vyhovuje zadání a doporučuji její uznání jako práce diplomové.

Petra Hoffmannová