BP

Úvod

**Úvod**

**Preco to robim, coho sa to tyka potivacia mojej prace urobit nieco koli niecomu**

Keďže sa riadenie letovej prevádzky (ATC) zaoberá poskytovaním pokynov a informácií o letovej prevádzke tak, aby sa zabezpečila bezpečná a plynulá prevádzka lietadiel. Cieľom tejto práce je, aby sa dalo s istotou identifikovať, akým smerom komunikácia prebieha. Kvôli veľkému širokospektrálnemu šumu v pozadí a zložitým termínom používaných v reči, v porovnaní s bežnou komunikáciou, totiž nie je vôbec jednoduché určiť smer komunikácie. Podľa publikácie Misunderstandings in ATC Communication, častokrát až spätný pohľad na komunikáciu vyjasní, že si ju interpretovali nesprávne. V reálnom čase by to znamenalo, že korekciu zle podanej informácie, je treba opraviť hneď a predísť tým prípadným nehodám ktoré by v rámci leteckej prepravy mohli mať katastrofálne následky.

Cieľom tejto práce a zároveň aj jej dôvodom by v konečnom následku malo byť zjednodušenie identifikácie zdroja ATC komunikácie a tým aj jednoduchšie rozpoznanie zdroja komunikácie pre osobu tretej strany. Ako prvý krok bolo obsiahnutie dát s ktorými s následne bude pracovať. K dispozícii som mal viacero nahrávok nepretržitej komunikácie z rôznych letísk a aj ich prepis do textového dokumentu.

Aby mohlo dôjsť k automatickému rozpoznávaniu, bolo najprv nutné rozdeliť danú nahrávku na samostatné súbory a takisto ich uložiť podľa smeru komunikácie spolu s vzostupným číslovaním pre prehľadnú analýzu podľa prepisu. Následne už rozdelené nahrávky boli analyzované pomocou skriptu v programovacom jazyku python a takisto boli extrahované dáta o amplitúde v čase, intenzite hlasu v čase a v neposlednom rade aj údaje o základnej frekvencii F0 v čase danej nahrávky. Po ich získaní boli dáta porovnané podľa toho, čo majú spoločné a čo odlišné v rámci rozdelenia podľa smeru komunikácie a na základe toho, bolo možné rozlíšiť, aké hodnoty údajov sprevádzajú komunikáciu v určitom smere.

**História a prehľad ATC**

Riadenie letovej prevádzky sa vzťahuje na „predchádzanie zrážok lietadiel a podporu usporiadaného toku letovej prevádzky“. Okrem toho sa vzdušný priestor, kde sa vykonáva riadenie letovej prevádzky, nazýva riadený vzdušný priestor (anglicky: Controlled airspace). Amplitúdová modulácia sa celosvetovo používa ako metóda modulácie rádiových vĺn pre rádiotelefóny používané na riadenie letovej prevádzky. Spočiatku, pri skorých začiatkoch letectva, nebola žiadna, verbálna ani neverbálna, komunikácia medzi pilotom a letiskom. Neskôr, sa však postupne schválil tzv. The Air Commerce Act, ktorý mal dopomôcť rýchlo sa rozširujúcej leteckej doprave prostredníctvom letísk a rôznych navigačných prvkov. Pred skorými 30. rokmi minulého storočia bola jediný bezpečnostný prvok, ktorým bol let cez deň a za priaznivých podmienok. Postupne sa ale bezpečnosť vyvíjala. Najskorší spôsob ATC bol človek, ktorý stál na viditeľnej pozícii letiska a mával farebnými vlajkami na letisku St. Louis, Missouri v 1929. Kvôli viacerým problémom, ako napríklad lety so zlou viditeľnosťou, sa postupne vlajky prestali používať a nahradili ich svetelné pištole, ktoré vedeli zasvietiť lúč svetla priamo na konkrétne lietadlo. Táto metóda sa využíva v takmer všetkých riadiacich vežiach dodnes ak náhodou nie je možné využiť rádiovú komunikáciu. Takisto sa od 30. rokov značne nezmenili ani signály. Keďže dopyt po leteckej preprave bol veľký, tak ani tento spôsob už nebol dostatočný na pomaly sa zhusťujúcu leteckú prepravu. Kvôli tomu bolo treba vyvinúť nejaký obojsmerný typ komunikácie, ktorým bola rádiová komunikácia. Počiatky rádiovej ATC komunikácie sú z letiska v Clevelande, Ohio. Práve tu bola postavená veža na starom hangári a boli v nej používané 15W vysielačky s dosahom až 15 míľ pri rôznych poveternostných podmienkach. Problém však bol s ich počiatočnou veľkosťou a hmotnosťou alebo, pri malých lietadlách, s nedostatkom energie na ich napájanie.

**Prozódia**

Reč je skladbou viacerých foném, ktoré sa následne spájajú do slabík a vytvoria slovo. Reč, ako ľudstvu najprirodzenejší komunikačný prostriedok, si teda rýchlo zabrala prvenstvo v ATC komunikácii. Zároveň je aj najefektívnejšia. Priemerne dokáže človek povedať 120 až 250 slov za minútu a tak niet pochýb, že v rámci hustej leteckej dopravy, je reč tým najlepším komunikačným kanálom. Prozódia sa vzťahuje na všetky fonetické vlastnosti, ktoré sa vyskytujú v reči. Zahŕňa intonáciu alebo výšku tónu, prízvuk, dôraz, relatívnu dĺžku tónu, rytmus, energiu, pauzy, kvalitu hlasu a intonáciu. Vzťahuje sa na tie prvky, ktoré sa môžu líšiť v závislosti od kontextu a dôrazu. Sú jazyky, kde je dôraz v určitom slove permanentne určený, no keďže v ATC komunikácii sa používa prevažne angličtina, musíme brať ohľad na to, že v angličtine má viacero rovnakých slov iný význam práve podľa toho, na ktorú slabiku dáme dôraz. Preto je aj v anglickej ATC komunikácii nesmierne dôležité, ktorým slovám dáme dôraz na aké miesto. Intonácia je ďalším dôležitým aspektom prozódie. Intonáciu vieme vyjadriť ako zmenu výšky hlasu v čase. Z toho sa dá v konečnom dôsledku odvodiť aj základná frekvencia tónu hlasu F0 v čase. Pre mužov je to zväčša medzi 80 až 200 Hz, pre ženy 180 až 400 Hz. Medzi týmito hodnotami F0  buď stúpa alebo klesá. Ďalšou prozodickou vlastnosťou je rytmická zmena. Rytmickou zmenou sa rozumejú rôzne dĺžky zvukov v slove. Slabika v slove je základnou jednotkou časovania v reči. Čo sa týka analýzy všetkých týchto vlastností reči, sú jedinečné u každého jedinca a tak sa stávajú akýmsi unikátnym identifikátorom človeka. Práca sa zameriava na konkrétnu analýzu týchto vlastností pochádzajúcich z viacerých zdrojov a ich A picture containing chart

Description automatically generatednásledné porovnanie, za účelom analýzy podobných a rozličných aspektov komunikácie.

Prozódia je taktiež podstatným faktorom v rámci rozpoznávania reči. Rozpoznávanie je hrubo založené na spektrálnych/kepstrálnych vlastnostiach získaných pomocou spektrálnej analýzy. Hlavným faktorom k dosiahnutiu očakávaných výsledkov je získanie hlasového záznamu v dobrých akustických podmienkach. Pri zlej akustike sa totiž v spektrálnej analýze reči vyskytuje oj chyba, ktorá vznikne zlými akustickými podmienkami, čo znamená že v analýze nebude len čistá reč ale aj napríklad ruch, ktorý vzniká akustikou miestnosti. Samotné spektrum hlasu hovoriacej osoby je teda značne rozdielne v akusticky ideálnej miestnosti, vo vonkajšom prostredí a taktiež v kokpite pilota alebo v samotnej riadiacej veži. Rovnako tento problém môžeme pozorovať pri ATC komunikácii, kde nám spektrum ovplyvňuje šum prenosového kanálu. Pri analýze ATC komunikácie teda musíme brať v úvahu všetky faktory ktoré ovplyvňujú spektrum a vo výsledkoch ich aj zohľadniť. Pôsobenie týchto faktorov si je možné všimnúť na ukážke nižšie pri vyslovení frázy “Don’t carry an oily rig like that“. Kanály pomenované TIMIT, NTIMIT a CTIMIT sú ukážkou variácie F0 pri nahrávaní blízko pri mikrofóne, cez kanál s vysokým obsahom šumu a v prostredí mobilnej siete. Pri charakteristikách kontúr F0 sa dá potvrdiť, že robustnosť prozodických vlastností je natoľko silná, že zmeny kontúr F0 sú minimálne, no stále Chart

Description automatically generatedpostrehnuteľné.

Letter

Description automatically generated with low confidence

Pri reči netreba zabúdať ani na emóciu. Rozdielne kontúry F0 môžeme získať z toho istého zdroja len pri emočnom rozdiele v reči.

**Text, letter

Description automatically generated**

Takisto emočný faktor vplýva aj na rozdiely tempa a páuz.

**Špecifické vlastnosti reči rečníka**

Každý rečník má svoje unikátne charakteristické črty, buď podľa rečníkových fyziologických charakteristík hlasových orgánov alebo nejakých zvykov. Rozdiely môžu byť zhrnuté nasledovne:

1. Veľkosť a tvar vokálneho traktu
2. Excitačné charakteristiky
3. Prozódia
4. Idiolekt
5. Sémantika

Prvé tri z vyššie uvedených čŕt vieme vyzistiť zo signálu a zvyšné dve sú prítomné v samotnom obsahu textu. Ukážka rozdielnosti rečníkov je zobrazená na obrázku ako rozdielne kontúry F0 pri rozprávaní čísloviek od 1 do 10 dvomi rozdielnymi ženami.A picture containing graphical user interface

Description automatically generated

**Pravdepodobnosť rozpoznania**

Rozpoznanie sa dá nazvať ako priradenie najpravdepodobnejšieho rozprávača, jazyka, emócie alebo nejakej zvukovej jednotky C\* vstupnej reči z množiny existujúcich rozprávačov, jazykov, emócii ale zvukových jednotiek. Ak je určené že označuje množinu existujúcich rozprávačov, jazykov, emócii ale zvukových jednotiek a O označuje pozorovania zistené zo vstupnej reči, tak rovnica rozpoznávanie v rámci pravdepodobnosti bude znieť nasledovne:

,

kde je aposteriórna pravdepodobnosť členu , pre daný výskyt reči vyjadrený pre .

**Jazykovo špecifické aspekty reči**

Je viacero rečových aspektov, ktoré sú rozdielne pri rozličných jazykoch ako napríklad akustická fonetika, prozódia, fonotaktika, slovník používaných slov a slovná skladba. Keďže ATC komunikácia pri lokálnych letoch neprebieha len v anglickom jazyku, tieto aspekty treba brať v úvahu aj pri analýze v rozličných jazykoch kde sa napríklad intonácia, rytmus alebo dôraz pri vyslovovaní rôznych slov mení v závislosti od hovoreného jazyka.

**Riešenie**

Cieľom semestrálnej práce, bol zber dát na samotnú analýzu z nahrávok nepretržitej ATC komunikácie medzi pilotmi a vežami z viacerých letísk.

Zvukové súbory, spolu s prepismi komunikácie, boli poskytnuté vedúcim práce. Zvukové súbory boli vo formáte .sph, čo je audio súbor, vytvorený vo formáte NIST SPHERE. Daný formát sa využíva práve v hlasovom rozpoznávaní a vlny v tomto formáte, majú bežne 16-bitovú pulzne kódovú moduláciu a vzorkovaciu frekvenciu o výške 16 KHz. Formát .sph však python neprečíta a tak bolo nutné danú nahrávku prekonvertovať do formát, s ktorým python problém mať nebude. To bolo docielené za použitia programu SoX, ktorý audio súbor prekonvertoval na Waveform Audio File Format (.wav).

Analýze samotných dát predchádzalo rozčlenenie nepretržitej, cca. dvojhodinovej komunikácie na segmenty. Jeden segment komunikácie sa rozumie ako jeden súbor s príponou .wav kde je extrahovaná práve jedna pasáž komunikačného kanála. Rozdelenie bolo dosiahnuté pomocou skriptu v programovacom jazyku python, kde v prvom kroku, je modifikovaný textový dokument s prepisom komunikácie, za účelom dosiahnutia formátu, ktorý obsahuje jedine údaj smeru komunikácie a počiatočný a konečný čas segmentu.

V druhom kroku, za použitia knižnice pydub, sa z modifikovaného textového dokumentu načítajú dáta a pomocou cyklu “for“, sa určí, že pri riadku, ktorý začína na PA (pilot veži) resp. AP (veža pilotovi), sa audio súbor segmentuje podľa časových údajov a následne ho uloží do príslušného priečinku podľa smeru komunikácie, rovnako do súboru s príponou .wav.

Po rozdelení segmentov do patričných priečinkov, bolo získaných vyše 1000 segmentov komunikácie z jednej nepretržitej. 1000 bola dostatočná vzorka na zistenie bližších detailov. Tie boli zistené pomocou pythonu v kombinácii s programom Praat. Praat je aplikácia, ktorá je určená na analýzu hovoreného slova a dokáže zo vzorky audio súboru analyzovať reč a následne vygeneruje konkrétne dáta, ktoré sa viažu s danou vzorkou reči. Pre analýzu ATC komunikácie sú obzvlášť potrebné dáta o intenzite a základnej frekvencii hlasu F0. Analýza samotných vzoriek bola obsiahnutá pomocou Praat knižnice pre python, parselmouth, tak, že cez cyklus for, boli postupne analyzované všetky segmenty komunikácie v obidvoch smeroch. Ku každému segmentu, boli priradené grafy analýz pomocou knižnice matplotlib. Grafy boli vygenerované pre amplitúdu hlasu v čase, intenzitu hlasu v čase a základnú frekvenciu F0 v čase pre každý konkrétny segment komunikácie. Pre porovnanie a určenie podobností a rozličností dát, boli taktiež vytvorené boxploty, pre porovnanie všetkých údajov jednej strany komunikácie. Na základe ich porovnania, je možné určiť, čo majú navzájom spoločné a odlišné a na základe toho, vytvoriť

**Záver**

Zadaním mojej práce, je pri kompletnom toku konverzácie medzi vežou a pilotom vedieť automaticky rozlíšiť a identifikovať, ktorá správa pochádza od konkrétneho zdroja.