

TransPraha: Hlasová navigace pro MHD

Petra MAREŠOVÁ, Jakub DOLEŽAL

Výzkumné a vývojové centrum pro mobilní aplikace
České vysoké učení technické v Praze, Fakulta elektrotechnická
Email: marespe5@fel.cvut.cz, jakub.dolezal@fel.cvut.cz

Článek se věnuje problematice samostatného pohybu zrakově handicapovaných v městské hromadné dopravě. Jako možné řešení je zde navržena, realizována a otestována hlasová navigace TransPraha, která umožňuje pomocí mobilního telefonu a systému, nasazeného na serverové straně, poskytnout informace o dostupném spojení i v situacích, kdy se uživatel dostane do zcela neznámého prostředí. Nalezené spoje jsou uživateli doporučeny s ohledem na obtížnost přestupů a časovou složitost. TransPraha disponuje hlasovou a hmatovou modalitou interakce, je ovládána pomocí kombinace tónové volby, rozpoznávání a syntézy řeči. Pro uživatele se zbytky zraku je navíc navržen klient pro chytré mobilní telefony, který nabízí jednoduché, kontrastní grafické rozhraní. Výsledky prvních testů s potenciálními uživateli potvrdily použitelnost aplikace. TransPraha bude během roku 2013 nasazena jako součást asistenční linky Navigačního centra Sjedené organizace nevidomých a slabozrakých.

Úvod

Samostatný pohyb osob je jedním z nejdůležitějších parametrů kvalitního života. Pro orientaci v prostoru využíváme především zrakové a sluchové vjemy. Díky tomu, že samotným zrakem získáváme až 80% informací [1], mají osoby se zrakovým handicapem velmi ztížené podmínky orientace. Běžné činnosti jako čtení jízdního řádu, nástup do tramvaje, přestup mezi stanovišti nebo zpětná kontrola správného spoje, které vidomý člověk řeší denně, automaticky a bez obtíží, se pro nevidomé mohou stát nepřekonatelnou překážkou. Tyto překážky se kumulují zejména v neznámém, entropickém prostředí, plném proměnlivých faktorů. Typickým příkladem je městská hromadná doprava (MHD).

Využití MHD při pohybu po městě je z finančního a časového hlediska pro nevidomé nezbytné. Ale závislost na cizí asistenci má dopad na všechny aspekty života handicapovaných osob. Nahrzení osobního asistenta vhodnou asistenční službou nebo zařízením je jedním ze způsobů, jak zrakově handicapovaným zprostředkovat potřebné informace pro podporu jejich samostatného pohybu a orientace při cestě MHD. Takové řešení je netriviální vzhledem k rozličnému charakteru potřeb uživatelů. Zahrnuje důkladnou analýzu principů prostorové orientace zrakově handicapovaných a integraci dostupných technologií, které uživateli zpřístupní relevantní informace natolik obsáhlé, aby se orientoval v prostředí bez asistence vidomých osob.

K vyřešení nastíněných problémů a ke zvýšení kvality cestování slabozrakých a nevidomých byla vyvinuta TransPraha – interaktivní, multimodální navigace pro MHD, která bude provozována jako součást asistenční linky Navigačního centra SONS (Sjedené organizace nevidomých a slabozrakých ČR)¹. TransPraha je primárně určena pro uživatele, kteří se samostatně pohybují v prostředí MHD. Pomáhá jim reagovat na změny cesty tím, že poskytuje informace o aktuálně dostupných dopravních spojení, a nabízí uživateli možnost zvolit si vyhovující variantu podle jeho preferencí.

Stav řešení problematiky

Informace o dopravním spojení získávají nevidomí majoritně z portálu Idos², který nabízí speciální webové rozhraní pro nevidomé tzv. „Blind friendly“ web. Toto řešení ovšem vyžaduje přístup na internet. V terénu je pak uživatel bez mobilního připojení při neočekávané situaci odkázán na pomoc vidomého kolemdoucího nebo asistenta na telefonní lince.

Díky integraci TTS (Text to Speech) se různé mobilní aplikace, nabízející vyhledání aktuálního dopravního spojení, staly přístupné i pro zrakově handicapované. Aplikace často nejsou psány primárně pro TTS, proto výstup TTS není ideální. Jako příklad můžeme uvést online aplikaci pro OS Android MHDroid³ nebo Pubtram⁴. Z aplikací nabízejících jízdní řády offline pod placenou licenci lze uvést CG Transit⁵, která je

¹ <http://navigace.sons.cz/>

² <http://blind.jizdnirady.idnes.cz>

³ <https://sites.google.com/site/androidmhdroid/>

⁴ <https://play.google.com/store/apps/details?id=cz.fhejl.pubtran>

k dispozici pro OS Andorid, platformu iOS i pro Windows Phone. Mezi multimodálními navigacemi zaujme adaptivní navigace Filipa Havlíčka na platformě Android [2], která propojuje navigaci ve volném prostoru s navigací v MHD a nabízí i rozhraní pro nevidomé uživatele, založené na principu čtyřrohového ovládání.

Často využívanou službou je telefonní asistenční linka NC SONS. Tu obsluhují specializovaní operátoři, kteří klientovi na počkání najdou aktuální spojení, doplní ho o další potřebné informace, dostupné na internetu, či poskytnou jednu z dalších služeb centra jako je například tvorba cestovních itinerářů či pomoc v nouzi. Nevýhodou jsou finanční nároky na provoz linky a fakt, že hovory nemohou být jedním operátorem vedeny paralelně. Operátoři obslouží 30 000 požadavků za rok, z toho tvoří 80% právě dotazy na spojení, které lze snadno automatizovat.

Popis navigace TransPražha

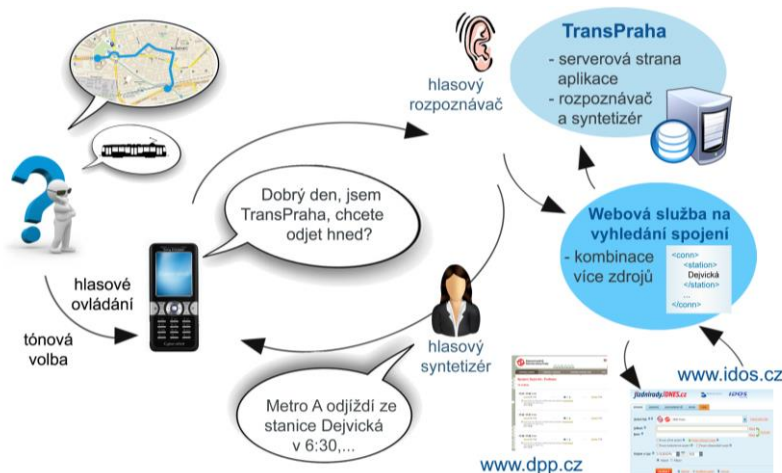
Pokud chce uživatel využít služby navigace TransPražha, zavolá obyčejným mobilním telefonem na příslušné číslo. K interakci se využívá rozpoznávání a syntéza řeči společně s tónovou volbou. Slabozrací uživatelé mohou využít klienta pro chytrý mobilní telefon (tzv. smartphone), který má navíc kontrastní grafického rozhraní.

Kromě informací o aktuálním dopravním spojení nabízí TransPražha další doplňující informace, které využijí právě zrakově handicapovaní uživatelé. Provádí analýzu spojů a na základě jejich obtížnosti, doby odjezdu a délky je nabídne uživateli v pořadí od nejsnazších k obtížnějším. TransPražha rovněž obsahuje algoritmus, který pomáhá řešit problém detekce více tramvají, přijíždějících na nástupiště ve stejný čas. V tomto případě může uživatel nezaregistrovat spoj, který zastaví až za soupravou stojící na začátku nástupiště. TransPražha nalezne spoje, které přijíždějí na nástupiště v přibližně stejný čas jako spoj, na který uživatel čeká, a informuje ho o nich.

Návrh navigace TransPražha

Abychom porozuměli potřebám zrakově handicapovaných při pohybu v prostředí městské hromadné dopravy, je nezbytné nejprve pochopit mechanismy jejich samostatného pohybu a prostorové orientace. Je třeba si uvědomit, že zrakově handicapované osoby nejsou homogenní skupina a informace, které potřebují k orientaci, mohou mít rozdílně obsáhlou informační hodnotu. Charakter potřebných dat je ovlivněn několika faktory. Patří mezi ně například míra a druh zrakového handicapu, povaha handicapované osoby, věk, ve kterém se zraková vada projevila, předchozí stav zraku či způsob výuky při absolvování školení prostorové orientace, který se může v rámci různých organizací lišit.

Celý proces návrhu s následnou realizací TransPražhy byl založen na diskusi a testování s potenciálními uživateli a se zástupci SONS. Na základě jejich odezvy byla pak TransPražha upravena a doplněna o další funkce. Způsoby interakce proto odrážejí požadavky uživatelů na obsah poskytnutých informací a způsob ovládání.



Obr. 1: Scénář interakce uživatele s TransPražha. Uživatel plánuje cestu hromadnou dopravou, zavolá tedy na linku TransPražha, obsluhovanou serverem. Pomocí rozpoznávání a tónové volby zadá své požadavky, pomocí syntézy řeči se dozvídá pokyny a informace o spojení.

⁵ <http://www.svetandroida.cz/cg-transit-jizdni-rady-v-kapse-i-bez-internetu-201209>

Uživatelský scénář

Uživatel může použít libovolný telefon k zavolání na asistenční linku SONS, viz Obr. 1. Při začátku hovoru se spustí dialog s TransPrahau, během kterého je postupně vyzván k zadání parametrů dopravního spojení. Následně server vyhledá dostupné spoje položením dotazu na webovou službu, např. IDOS⁶, a provede jejich analýzu na základě jejich časové složitosti, počtu, obtížnosti a fyzické délce přestupů a délce celého spojení. Je možná podpora více webových služeb na vyhledání spojení pro pokrytí různých měst i redundanci pro případ výpadku. Nakonec jsou spoje prezentovány uživateli formou syntézy řeči TTS.

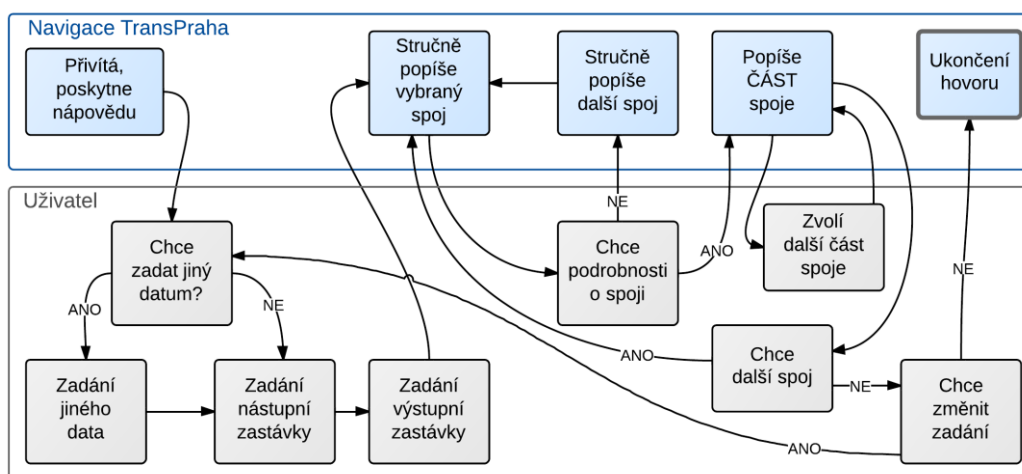
TransPraha podporuje alternativní variantu scénáře, určeného pro mobilní telefon s detekcí GPS polohy uživatele. Klient před začátkem hovoru odešle polohu na server, který na jejím základě uživateli nabídne blízké nástupní zastávky.

Rozhraní telefonního hovoru

Uživatel komunikuje pomocí telefonního hovoru se serverem, spravovaným SONS. V rámci hovoru je přenášén od uživatele akustický signál řeči či tónové volby, která na rozdíl od řeči umožňuje ovládat aplikaci i v hlučném prostředí. Server na základě získaných informací od uživatele vytvoří odpověď a syntetizuje ji. Dialog mezi serverem a uživatelem se skládá z řady propojených stavů, zobrazených na Obr. 2, ve kterých uživatel postupně zadává své požadavky. Průchod stavy může být dynamicky ovlivněn například vyvoláním nápovědy, kterou uživatel spustí takzvaným univerzálním příkazem (např. „Porad’ mi“).

Obr. 2 popisuje střídání iniciativy mezi uživatelem a serverem a přechody mezi stavy dialogu. Po přivítání uživatel zadá v několika následujících stavech potřebné informace, na základě kterých je poté seznámen se všemi vybranými spoji. Pro každý spoj je z důvodů úspornosti syntetizováno malé množství nejdůležitějších informací, např. odjezd, doba trvání a počet přestupů. Uživatel si na jejich základě může vybrat preferovaný spoj a je mu syntetizován jeho detailní popis. Jednotlivé části jsou pak uživateli syntetizovány po částech. Po syntéze jedné části je vyzván k zadání povelu „další“, který spustí syntézu následující části.

První stav dialogu uživatele přivítá a upozorní ho na možnost nápovědy. Při přechodu do dalšího stavu jsou uživateli syntetizovány potřebné informace, popřípadě je vyzván, aby zadal svůj požadavek. Rozpoznávání na straně serveru nerozpoznává kontinuálně, ale porovná vstup uživatele s předem definovanou skupinou očekávaných povelů, tzv. gramatikou (například gramatika obsahující pouze výrazy „ano“, „ne“, či gramatika obsahující všechny názvy dostupných zastávek). Pokud uživatel mlčí, či řekne něco neočekávaného, je opakovaně naveden ke správné odpovědi kontextovou nápovědou. Pokud selže rozpoznávání z důvodu hlučného prostředí, je vyzván k použití tónové volby⁷.



Obr. 2: Schéma interakce mezi uživatelem a Navigací během telefonního hovoru. Obdélník značí stav: šedé stavy vyžadují volbu uživatele, modré stavy syntetizují informace, např. informace o spojení. Kdykoliv během dialogu lze použít univerzální řečové příkazy jako například „porad’ mi“.

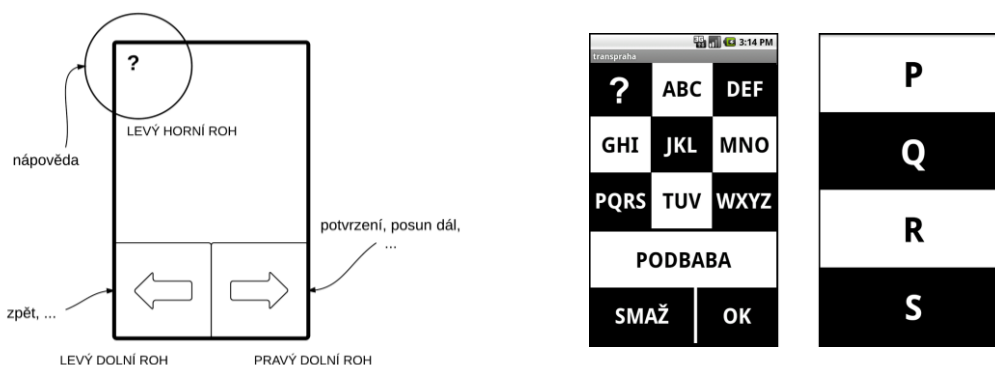
⁶ <http://jizdnirady.idnes.cz/praha/spojeni/>

⁷ Dual-tone multi-frequency signaling (DTMF)

Klient pro chytré mobilní telefony

Telefony disponující hardwarovou klávesnicí a OS Symbian pomalu mizí z trhu a jsou nahrazovány smartphony s dotykovým displejem. V rámci projektu byl navržen a vytvořen prototyp klienta pro telefony s dotykovým displejem a lokalizátorem polohy GPS. Ten je určen pro uživatele se zbytky zraku. Klient je ovládán přes grafické rozhraní (GUI). V současné době se pracuje na propojení grafického rozhraní s hlasovým, aby klient nabízel souběžné ovládání pomocí více modalit. GUI je vysoce kontrastní. Používá v celém rozsahu aplikace stejný princip ovládání, popsany na obrázku Obr. 3 vlevo. Ten využívá primárně dobře nahmatatelných rohů displeje. Levý horní roh vždy vyvolá kontextovou nápovědu. Levý spodní roh je používán k vrácení akce, posunu zpět či mazání, pravý roh zase k posunu dál či potvrzení. Střed displeje je využit pro libovolně přibližitelný text, který stručně popisuje stav aplikace či konkrétní informace o spojení.

Výjimku tvoří speciální, jednoduchá klávesnice, která je zobrazena na Obr. 3 vlevo. Ta ve středu displeje obsahuje tlačítka, jejichž rozložení je inspirováno klasickou hardwarovou klávesnicí. Stiskem příslušné softwarové klávesy se zobrazí zvětšená část klávesy, rozdělená na 3-4 tlačítka, která reprezentují jednotlivá písmena. Každé písmeno je tedy možné napsat pomocí dvou stisků.



Obr. 3: Návrh kontrastního grafického rozhraní klienta. Nalevo je znázorněn princip rozložení ovládacích prvků, shodný pro všechna dialogová okna. V levé části je zobrazena dotyková klávesnice, umožňující zadat název stanice.

Ověření návrhu testováním s uživateli

Metodika testování

Důležitou součástí UCD (User centered design) je zapojení koncového uživatele už v raných fázích návrhu. To vede k tvorbě uživatelsky příjemného rozhraní a umožňuje to brzké odhalení a řešení chyb návrhu. Byla zvolena metoda testování použitelnosti (Usability testing). Použitelné rozhraní systému se lze rychle naučit ovládat a jeho princip je konzistentní v celém rozsahu systému. Testování bylo zaměřeno na použitelnost hlasového a hmatového rozhraní TransPraha a získání zpětné vazby od potenciálních uživatelů. Hlavním cílem bylo testování návrhu jednotlivých částí rozhovoru a systému kontextové nápovědy.

Za těmito účely byly zkoumány tyto parametry rozhraní: naučitelnost neboli schopnost naučit se ovládat daný systém a princip ovládání si zapamatovat, výkonnost (efektivitu), neboli schopnost snadno, rychle a úspěšně provést zadané úlohy a spokojenost uživatelů s rozhraním. Parametry použitelnosti byly vyhodnoceny na základě metrik, naměřených v průběhu testu. Byly sledovány následující metriky:

1. **Čas prodlevy reakce uživatele** na otázku (doba, kterou uživatel přemýšlí nad formulací odpovědi).
2. **Počet vyslechnutých nápověd pro každou otázku**, které se syntetizují v případě špatné odpovědi.

Každá metrika se zaznamenávala zvlášť pro jednotlivé gramatiky, neboli části rozhovoru, kde se očekává odpověď uživatele. Po testu uživatelé ještě zodpověděli dotazník, ve kterém subjektivně ohodnotili různé aspekty interakce s aplikací. V rámci experimentu byly ověřovány tři hypotézy:

- **hypotéza 1:** Rozhraní je naučitelné, prodleva se bude při dalším použití podobné odpovědi snižovat,
- **hypotéza 2:** Průměrná časová prodleva nepřesáhne stanovený limit 3 vteřiny,
- **hypotéza 3:** Průměrné subjektivní hodnocení každé zkoumané vlastnosti nebude nižší než 7.

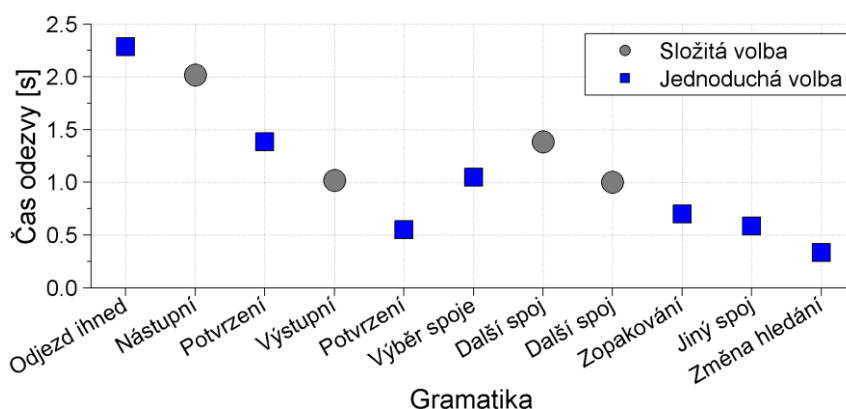
Hypotéza 1 a 2 byly ověřeny na základě rozboru první metriky. Hypotéza tři byla vyhodnocena podle výsledků dotazníků, které byly vyplněny s uživateli po závěru testování.

Průběh testu

Podle předběžných dotazníků bylo vybráno šest uživatelů různých charakteristik a zrakových handicapů. Uživatelé prošli dva scénáře, které popisovaly typické použití TransPrahy. V prvním scénáři hledali uživatelé aktuální spojení ze stanice Chodov do stanice Můstek. V druhém scénáři zkusili zadat jiné datum a jiný čas. Poslední scénář byl zaměřen na ověření samotného hmatového rozhraní, které poskytuje plnohodnotné ovládání aplikace v případě, že uživatel nemůže nebo nechce použít hlasové rozhraní. Uživatelé v něm prošli první scénář, ale byl jim vypnut mikrofon. Před testem záměrně neproběhla žádná instruktáž o způsobu ovládání aplikace, aby testování odhalilo nejednoznačné a neintuitivní části rozhraní.

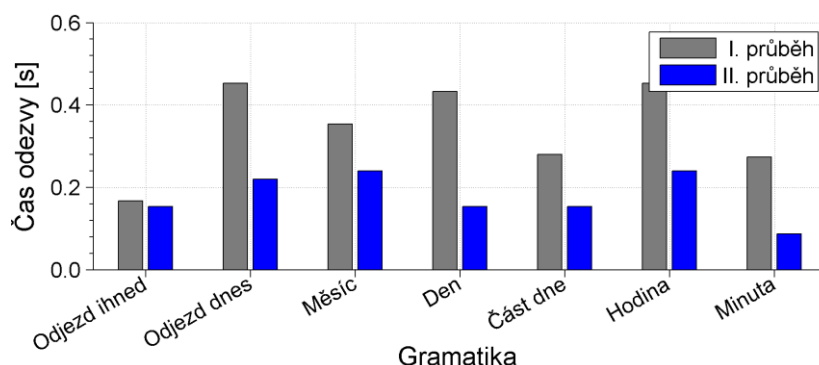
Diskuze výsledků

První metrikou byl čas prodlevy reakce uživatele na otázku. Na Obr. 4 je vidět, že prodleva u otázek typu „ano“/„ne“ má klesající tendenci. Výjimku tvoří výběr spoje, kde uživatel musel kromě formulace odpovědi přemýšlet i nad výběrem popsaného spoje. Za zmínku stojí i zadání a potvrzení nástupní stanice, které je následováno zadáním a potvrzením stanice výstupní. Zde je vidět, že uživatelé pochopili princip zadání nástupní stanice a aplikovali ho při zadání výstupní stanice, kde je výrazně menší zpoždění odpovědi.



Obr. 4: Prodlevy reakce uživatele pro jednotlivé otázky z prvního scénáře v pořadí v jakém dialog s aplikací probíhal. Modré čtverce označují otázky, na které se očekává odpověď „ano“/„ne“. Šedé kruhy označují otázky, na které se očekává složitější odpověď, jako je název zastávky nebo pokyn „další“.

Druhý scénář uživatelé prošli dvakrát, protože otázky naváděly na odlišnou formulaci odpovědi. Bylo sledováno zlepšení, které je zobrazeno na Obr. 5. Z něj je patrné, že největší problém dělala uživatelům formulace konkrétního dne a hodiny.



Obr. 5: Prodleva reakce při prvním a druhém průchodu scénářem 2.

První hypotéza, která předpokládala snižování času odezvy v průběhu interakce, byla potvrzena. Prodleva dokonce nepřesáhla tři vteřiny, což je v hlasových aplikacích maximální akceptovatelná doba bez interakce [3], čímž se potvrdila i druhá hypotéza. Znamená to, že uživatelé interakci s aplikací pochopili její ovládání i bez předchozí instruktáže a přípravy. Princip interakce si pamatovali a byli schopni ho použít v dalších částech dialogu.

Druhou metrikou byl počet vyslechnutých náповěd, které následovaly po špatné odpovědi. Rozhraní lze považovat za efektivní díky tomu, že byli uživatelé schopni splnit 100% úloh scénáře, a to s relativně malým počtem náповěd. Ty uživatelé potřebovali ke správné formulaci odpovědi v prvním scénáři průměrně v jednom případě z jedenácti. V druhém chybovali průměrně v jedné odpovědi ze sedmi. Nejčastěji při zadání hodiny, kde tři ze sedmi účastníků využili první náповěd. Počet špatných odpovědi na otázku poukazoval na to, že je nejasně formulovaná otázka nebo nedostatečně definovaná skupina přijatelných odpovědi. Gramatiky, kde byl zaznamenán nejčetnější výskyt stejných chyb (chyboval více než jeden uživatel), byly přeformulovány. U třetího scénáře zaměřeného na hmatové rozhraní splnilo zadání pět uživatelů z šesti. Jeden účastník měl problémy se zadáním názvu zastávky, který kopíruje styl psaní SMS, protože SMS ze svého mobilního telefonu neposílá.

Protože návrh sleduje principy metodiky User centered designu (UCD), byly subjektivní dojmy a názory testujících osob na interakci s aplikací velmi důležité. Přehlednost, snadné použití, intuitivnost ovládání a další parametry hodnotili uživatelé v dotazníku známkami v intervalu osm z deseti až devět z deseti, což potvrdilo i třetí hypotézu. Reakce na ovládání byly během testování kladné, což se projevilo i na výsledném hodnocení.

Závěr

TransPraha – hlasová navigace pro MHD – byla navržena a realizována v podobě funkčního prototypu. Byla ověřena funkčnost ovládání pomocí rozpoznávání a syntézy řeči, tónové volby. Algoritmy TransPrahy dolují potřebná data z webu a rozšiřují popis spojení o další informace, které jsou určeny pro zrakově handicapované uživatele. TransPraha částečně pokrývá speciální potřeby uživatelů: 1. varianty řečových dialogů dle pokročilosti uživatele, 2. upřednostnění spojů bez komplikovaných přestupů, 3. upozornění na více linek, příjždějících do nástupní stanice. Dle provedených uživatelských testů jsou slabozrací a nevidomí uživatelé schopni se rychle naučit s navigací pracovat a ovládání si zapamatovat.

Následující část vývoje a testování projektu je zaměřena na nasazení TransPrahy v rámci asistenční linky NC SONS. Je plánováno rozšíření funkcionality o sestavení cestovního itineráře mezi libovolnými místy. Pro mobilní telefony s detekcí GPS polohy bude navigace pro MHD propojena s navigací ve volném prostoru. Bude tak navádět uživatele z jeho současné polohy na konkrétní nástupiště nejbližší stanice, přes MHD a z výstupní zastávky cílové stanice až do cíle. Dále je v plánu rozšířit prototyp klienta pro kombinaci grafického i řečového ovládání. Informace o spojení pak budou dostupné zároveň v audio i vizuální podobě. Pro nevidomé uživatele bude přínosná funkce dostupnosti audio verze informací i v případě odpojení od linky.

Poděkování

Projekt TransPraha je realizován ve spolupráci se Sjedenou organizací nevidomých a slabozrakých a financován Nadací Vodafone Česká Republika.

Literatura

1. Wiener, P.: Prostorová orientace zrakově postižených. Institut rehabilitace zrakově postižených UK FHS, třetí vydání, Praha, 2006.
2. Havlíček, F.: Adaptive Multimodal Navigation for Android. Czech Technical University in Prague, Faculty of Electrical Engineering, Prague, 2011. http://dip.felk.cvut.cz/browse/pdfcache/havlifi4_2011dipl.pdf
3. Harris, R. A.: Voice Interaction Design: Crafting the New Conversational Speech Systems, Morgan Kaufmann, 2004.