

Teme za studentske radove iz oblasti: "Razvoj aplikacija govornih tehnologija i/ili dijaloških sistema za pametne uređaje"

Teme:

1. Razvoj dijaloških strategija i unapređenje dijaloških sistema (pametna kuća sa upravljanjem glasom za pametni TV, tablet računar, telefon i/ili sat, robot MARKO; govorni automati za informacije o sportskim rezultatima, redu vožnje, vremeskoj prognozi, itd...)
2. Pomagala za decu, stare i/ili osobe sa invaliditetom, bazirana na primeni govornih tehnologija
3. Aplikacije sinteze govora iz teksta (ozvučavanje sajtova, audio biblioteke za studente, itd...)
4. Razvoj ekspresivnosti sinteze govora na osnovu teksta (npr. dodavanje emocija u neutralni govor)
5. Implementacija TTS-a na Andoid OS (cilj npr. PDF/Word reader na telefonima i tablet računarima)
6. Sintetizovanje i kodovanje govora na bazi modelovanja artikulacije govora
7. Animacija lica koje govori (npr. TalkingHead na srpskom, razvoj avatara u društvenim aplikacijama)
8. Audio-vizuelno prepoznavanje govora i emocija (npr. prepoznavanje emocija korišćenjem kamere i mikrofona smart uređaja (mobilni telefon, tablet računar, TV))
9. Detekcija govora pod stresom i adaptacija automatskog prepoznavanja govora
10. Automatsko i forenzičko prepoznavanje govornika (npr. zaštita smart uređaja putem glasa)

Oblast i motivacija:

Govorne tehnologije omogućuju govornu komunikaciju sa pametnim uređajima. U jednom smeru vrši se sintetizovanje govora na osnovu teksta, a u drugom automatsko prepoznavanje govora, govornika, pa i emocija. U dijalogu čovek-mašina, govorne tehnologije se naslanjaju na jezičke tehnologije, tj. obradu prirodnog jezika, kao i modul za upravljanje dijalogom koji vrši kognitivne funkcije dijaloškog sistema.

Projekat tehnološkog razvoja "*Razvoj dijaloških sistema za srpski i druge južnoslovenske jezike*" (TR32035, 2011-15) učvrstio je FTN na poziciji lidera u razvoju i primeni govornih i jezičkih tehnologija u regionu. Pored publikacija u naučnim časopisima, rezultati ovog projekta su osnova za brojna tehnička rešenja (<http://www.ftn.uns.ac.rs/Tehnicka-Resenja-M80>) od kojih neka imaju široke primene i perspektive na tržištu, a neka imaju veliki društveni značaj kao što su pomagala za osobe sa invaliditetom.

Samsungova laboratorija za razvoj aplikacija je nedavno otvorena na FTN sa ciljem da studentima pruži priliku da razvijaju aplikacije za pametne uređaje najnovije generacije: pametne table i televizore, tablet računare, pametne telefone i satove.

Cilj ove grupe diplomskih radova je da se razvijaju aplikacije govornih tehnologija i dijaloških sistema koje omogućuju da se sa pametnim uređajima komunicira na srpskom i srodnim južnoslovenskim jezicima.

Ciljevi – aktivnosti:

- Analiza i evaluacija postojećih aplikacija govornih tehnologija i/ili dijaloških sistema.
- Specifikacija odgovarajućeg niza aktivnosti, razvoja komponenti i interfejsa, kao i načina testiranja.
- Implementacija unapređenih tehničkih rešenja aplikacija govornih tehnologija i/ili dijaloških sistema.
- Evaluacija unapređenih komponenti i aplikacija govornih tehnologija i/ili dijaloških sistema.

Potrebno predznanje:

- Analiza i obrada audio i/ili video signala.
- Elementarno poznavanje govornih tehnologija i/ili obrade prirodnog jezika.
- Programiranje: Java i/ili C++; Operativni sistemi: Android i Windows.

Reference:

Dijaloški sistemi i aplikacije govornih tehnologija (teme 1, 2 i 3):

- [1] "Challenges of Natural Language Communication with Machines" (2013) Delić V., Sečujski M., Jakovljević N., Gnjatović M., Stanković I. *Chapter 19 in DAAAM Int. Scientific Book 2013*, Eds: Katalinic B., Tekic Z., Vienna, Austria, pp. 371-388, DOI: 10.2507/daaam.scibook.2013.19
- [2] "How Speech Technologies Can Help People with Disabilities" (2014) Delić V., Sečujski M., Vujnović Sedlar N., Mišković D., Mak R., Bojanić M. *Proc. of the 16th Int. Conf. on Speech and Computer – SPECOM*, Novi Sad, Serbia, Eds: Ronzhin A. et al., Springer LNAI 8773, pp. 243-250, DOI: 10.1007/978-3-319-11581-8_30
- [3] "Therapist-Centered Design of a Robot's Dialogue Behavior" (2014) Gnjatović M. *Cognitive Computation*, Vol. 6, No. 4, pp. 775-788, DOI: 10.1007/s12559-014-9272-1
- [4] "Changing Concepts of Machine Dialogue Management" (2014) Gnjatović M. *Proc. of the 5th IEEE Int. Conf. on Cognitive Infocommunications*, Vietri sul Mare, Italy, pp. 367-372, DOI: 10.1109/CogInfoCom.2014.7020480
- [5] "Cognitively-inspired representational approach to meaning in machine dialogue" (2014) Gnjatović M., Delić V. *Knowledge-Based Systems*, Elsevier, Vol. 71, pp. 25-33, DOI: 10.1016/j.knosys.2014.05.001
- [6] "Focus Tree: Modeling Attentional Information in Task-Oriented Human-Machine Interaction" (2012) Gnjatović M., Janev M., Delić V. *Applied Intelligence*, Vol. 37, No. 3, pp. 305-320, DOI: 10.1007/s10489-011-0329-5
- [7] Tehničko rešenje: *Sistem za ozvučavanje Internet prezentacija* (2010) Delić V. i dr. FTN i AlfaNum, Novi Sad <http://www.ftn.uns.ac.rs/M81-Ozvucavanje-Internet-Prezentacija>
- [8] Tehničko rešenje: *Govorni softver za slepe i slabovide – anReader* (2010) Delić V. i dr. FTN i AlfaNum, Novi Sad <http://www.ftn.uns.ac.rs/M81-anReader>
- [9] Tehničko rešenje: *Sistem za red vožnje* (2012) Knežević D. i dr. FTN i AlfaNum, Novi Sad <http://www.ftn.uns.ac.rs/M81-Sistem-za-Red-Voznje>
- [10] Tehničko rešenje: *Audio biblioteka za slepe i slabovide* (2010) Mišković D. i dr. FTN i AlfaNum, Novi Sad <http://www.ftn.uns.ac.rs/M81-anABSS>
- [11] Tehničko rešenje: *Pametna kuća bazirana na govornim tehnologijama za srpski jezik – anSmartHome* (2014) Ostrogonac S. i dr. FTN i AlfaNum, Novi Sad <http://www.ftn.uns.ac.rs/M85-anSmartHome>
- [12] Tehničko rešenje: *Igra za slepa i slabovida lica bazirana na govornim tehnologijama za srpski jezik – anMasterMind* (2014) Ostrogonac S. i dr. FTN i AlfaNum, Novi Sad <http://www.ftn.uns.ac.rs/M85-anMasterMind>
- [13] Tehničko rešenje: *Sistem za davanje sportskih rezultata* (2012) Suzić S. i dr. FTN i AlfaNum, Novi Sad <http://www.ftn.uns.ac.rs/M85-Davanje-Sportskih-Rezultata>
- [14] Tehničko rešenje: *Robotski sistem sa integrisanim sistemima za obradu slike i govora i konverzionim agentom* (2012) Tasevski J. i dr. FTN, Novi Sad <http://www.ftn.uns.ac.rs/M85-Robot-MARKO>

Sinteza govora na osnovu teksta (teme 4, 5, 6 i 7):

- [1] "HMM-based Speech Synthesis for the Serbian Language" (2012) Pakoci E., Mak R. *Prof. of the 56th Conf. for ETRAN*, Zlatibor, Serbia, pp. TE4.7-1-4
- [2] "Automatska konverzija tekstualnih informacija u govor" (2011) Sečujski M., Delić V. Monografska serija, Naučnotehničke informacije, ISBN: 978-86-81123-25-6, ISSN 1820-3418, Vol. XLVI, No. 4, Vojnotehnički institut, Beograd, 56 strana
- [3] "Automatic Prosody Generation for Serbo-Croatian Speech Synthesis Based on Regression Trees" (2011) Sečujski M., Jakovljević N., Pekar D. *Proc. of the 12th Annual Conference of the ISCA – INTERSPEECH*, Firenze, Italy, pp. 3157-3160
- [4] "A comparison of decision tree based approaches to the modelling of prosodic features in Serbian" (2013) Sečujski M., Pekar D., Mak R., Pakoci E., Jakovljević N., Delić V., Popović B. In *Verbal Communication Quality, Interdisciplinary Research II*, Eds: S.T. Jovićić et al., Belgrade, Serbia, pp. 75-90
- [5] *Developments in Speech Synthesis* (2005) Tatham M., Morton K. ISBN: 978-0-470-85538-6, Wiley.
- [6] *Text-to-Speech Synthesis* (2009) Taylor P. ISBN: 978-0-521-89927-7, Cambridge Univ. Press.
- [7] "The HMM-based speech synthesis system (HTS) version 2.0" (2007) Zen H., Nose T., Yamagishi J., Sako S., Masuko T., Black A., Tokuda K. *Proc. of the 6th ISCA Workshop on Speech Synthesis*, pp. 294-299
- [8] Tehničko rešenje: *Sintetizator govora na bazi skrivenih Markovljevih modela za srpski jezik* (2012) Pakoci E. i dr. FTN i AlfaNum, Novi Sad <http://www.ftn.uns.ac.rs/M85-anTTS-HMM>
- [9] Tehničko rešenje: *Tehnologija automatske sinteze govora na osnovu teksta na srpskom i drugim južno-slovenskim jezicima* (2010) Sečujski M. i dr. FTN i AlfaNum, Novi Sad <http://www.ftn.uns.ac.rs/M81-anTTS>

Prepoznavanje govora, govornika i emocija (teme 8, 9 i 10):

- [1] "Discrimination Capability of Prosodic and Spectral Features for Emotional Speech Recognition" (2012) Delić V., Bojanić M., Gnjatović M., Sečujski M., Jovičić S. *Elektronika ir Elektrotehnika*, Vol. 18, No. 9, pp. 51-54, DOI: 10.5755/j01.eee.18.9.2806
- [2] "Discriminative Training for Automatic Speech Recognition" (2012) Heigold G., Ney H., Schluter R., Wiesler S. *IEEE Signal Processing Magazine*, Vol. 29, No. 6, pp. 58-69, DOI: 10.1109/MSP.2012.2197232
- [3] "Comparison of Linear Discriminant Analysis Approaches in Automatic Speech Recognition" (2013) Jakovljević N., Mišković D., Janev M., Sečujski M., Delić V. *Elektronika ir Elektrotehnika*, Vol. 19, No. 7, pp. 76-79, DOI: 10.5755/j01.eee.19.7.5167
- [4] "Speech Recognition with Weighted Finite-State Transducers" (2008) Mohri M., Pereira F., Riley M. *Chapter 28 in Springer Handbook of Speech Processing*, pp. 559-584, DOI: 10.1007/978-3-540-49127-9_28
- [5] "Speaker Detection Using Phoneme Specific Hidden Markov Models" (2014) Pakoci E., Jakovljević N., Popović B., Mišković D., Pekar D. *Proc. of the 16th Int. Conf. on Speech and Computer – SPECOM*, Novi Sad, Serbia, Eds: Ronzhin A. et al., Springer LNAI 8773, pp. 410-417, DOI: 10.1007/978-3-319-11581-8_51
- [6] "Large-Vocabulary Continuous Speech Recognition Systems: A Look at Some Recent Advances" (2012) Saon G., Chien J. *IEEE Signal Processing Magazine*, Vol. 29, No. 6, pp. 18-33, DOI: 10.1109/MSP.2012.2197156
- [7] **Tehničko rešenje: Tehnologija automatskog prepoznavanja govora na srpskom i njemu srodnim jezicima** (2010) Delić V. i dr. FTN i AlfaNum, Novi Sad <http://www.ftn.uns.ac.rs/M81-anASR>
- [8] **Tehničko rešenje: Dekoder za prepoznavanje kontinualnog govora na velikim rečnicima** (2012) Jakovljević N. i dr. FTN i AlfaNum, Novi Sad <http://www.ftn.uns.ac.rs/M85-anLVCSR-dekoder>

Kontakti:

- Dr Vlado Delić, red. prof, vdelic@uns.ac.rs
- Dr Milan Sečujski, docent, secujski@uns.ac.rs
- Dr Nikša Jakovljević, docent, jakovnik@uns.ac.rs
- Dr Milan Gnjatović, docent, gnjatovic@uns.ac.rs
- Dr Branislav Popović, istraživač-saradnik, bpopovic@uns.ac.rs
- Dr Milana Bojanić, istraživač-saradnik, milana.bojanic@uns.ac.rs

Nastavnici i saradnici sa projekta TR32035 zaduženi za teme studentskih radova:

Prof. dr Vlado Delić (uz asistenciju dr Milana Gnjatovića, Siniše Suzića i Dragiše Miškovića):

1. Razvoj dijaloških strategija i unapređenje dijaloških sistema (pametna kuća sa upravljanjem glasom za pametni TV, tablet računar, telefon i/ili sat; robot MARKO; govorni automati za informacije o sportskim rezultatima, redu vožnje, vremeskoj prognozi, itd...)
2. Pomagala za decu, stare i/ili osobe sa invaliditetom, bazirana na primeni govornih tehnologija
3. Aplikacije sinteze govora iz teksta (ozvučavanje sajtova, audio biblioteke za studente, itd...)

Doc. dr Milan Sečujski (uz asistenciju Roberta Maka na prve 3 teme i dr Branislava Popovića na 4. temi):

4. Razvoj ekspresivnosti sinteze govora na osnovu teksta (npr. dodavanje emocija u neutralni govor)
5. Implementacija TTS-a na Andoid OS (cilj npr. PDF/Word reader na telefonima i tablet računarima)
6. Sintetizovanje i kodovanje govora na bazi modelovanja artikulacije govora
7. Animacija lica koje govori (npr. TalkingHead na srpskom, razvoj avatara u društvenim aplikacijama)

Doc. dr Nikša Jakovljević (uz asistenciju dr Branislava Popovića, dr Milane Bojanić i Edvina Pakocija):

8. Audio-vizuelno prepoznavanje govora i emocija (npr. prepoznavanje emocija korišćenjem kamere i mikrofona smart uređaja (mobilni telefon, tablet računar, TV))
9. Detekcija govora pod stresom i adaptacija automatskog prepoznavanja govora
10. Automatsko i forenzičko prepoznavanje govornika (npr. zaštita smart uređaja putem glasa)