

Nastavno-naučnom veću
Fakulteta organizacionih nauka
Univerziteta u Beogradu

Predmet: Podobnost teme i kandidata Olje Latinović za izradu doktorske disertacije

Odlukom 3/89-13 br. od 15.10.2014. godine imenovani smo za članove Komisije za ocenu teme i podobnosti kandidata Olje Latinović za izradu doktorske disertacije i naučne zasnovanosti teme „Uticaj varijacija u interakciji korisnika na performanse sistema za prepoznavanje govornika“.

Na osnovu materijala priloženog uz Zahtev kandidata, Komisija podnosi sledeći

I Z V E Š T A J

1. Podaci o kandidatu

1.1. Biografski podaci

Olja Latinović rođena je 24. aprila 1988. godine u Prijedoru. Nakon završene gimnazije u Prijedoru (2006. godine) upisuje redovne studije na Panevropskom univerzitetu „Apeiron“ u Banja Luci – smer Inženjering informacionih tehnologija. Odličnom ocenom na odbrani diplomskog rada „MS SQL baza bibliotečkog informacionog sistema sa interfejsom u C#“, te 2010. postaje diplomirani inženjer informacionih tehnologija. Nakon završetka osnovnih studija u Banja Luci, upisuje na Fakultetu organizacionih nauka Univerziteta u Beogradu diplomatske akademske - master studije, studijski program Informacioni sistemi i tehnologije, i sa ocenom deset odbranom master rada stiče zvanje mastera. Doktorske studije na Fakultetu organizacionih nauka upisuje, studijski program Informacioni sistemi i kvantitativni menadžment, izborno područje Informacioni sistemi upisala je 2012.g.

Od 2009. do 2013. godine učestvuje u nastavi kao stručni saradnik za predmete Baze podataka, Projektovanje informacionih sistema, Poslovne aplikacije, Informacione tehnologije, te uskoro postaje asistent, zatim i viši asistent na Panevropskom univerzitetu „Apeiron“. Pored toga, učestvuje u organizaciji međunarodnog naučno-stručnog skupa Informacione tehnologije u elektronskom obrazovanju (ITeO) koji se svake godine održava u Banja Luci, te u radu seminara „Oblici i metode interaktivne nastave sa studentima“. 2013. godine počinje da radi u softverskoj firmi „Breza software engineering“ u Beogradu, kao softverski inženjer, gde je angažovana na realizaciji više projekata. U međuvremenu objavljuje naučne radove, najviše u sferi biometrije, sa saradnicima, asistentima i profesorima iz multimedijalne laboratorije sa Fakulteta organizacionih nauka. Pored toga, stiče zvanične „Oracle“ sertifikate kao Oracle Database 11g Performance Tuning Certified Expert, te Oracle Database 11g Administrator Certified Professional, i Oracle Certified Professional Java SE 7 Programmer. Govori srpski, engleski, nemački i francuski jezik.

Zaposlenje:

2010 – 2011 : Saradnik u nastavi, Panevropski univerzitet Apeiron, Banja Luka, BiH
2011 – 2012: Asistent, Panevropski univerzitet Apeiron, Banja Luka, BiH
2012 – 2013: Viši asistent, Panevropski univerzitet Apeiron, Banja Luka, BiH
2013 – trenutno: Softver inženjer, Breza Software Engineering, Beograd, Srbija

1.2. Stečeno naučnoistraživačko iskustvo

Za vreme diplomskih i doktorskih akademskih studija učestvovala je kao student saradnik u radu Laboratorije za multimedijalne komunikacije na projektu "Primena multimodalne biometrije u menadžmentu identiteta", TR32013, Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije.

Kandidatkinja je 2012. odbranila master rad, iz oblasti projekta TR32013, pod naslovom "Primena rešenja tehnologije otvorenog koda u prepoznavanju osobe zasnovane na glasovnom zapisu".

Publikacije:

Rad objavljen u časopisu nacionalnog značaja (M50)

1. Ivan Milenković, Olja Latinović, Dejan Simić, *Using Kerberos protocol for Single Sign-On in Identity management systems*, JITA – Journal of Information Technology and Applications, Banja Luka, B&H, Vol. 3, No.1, 2013. pp. 27-33

Rad objavljen na nacionalnom naučnom skupu (M60)

2. Olja Latinović, Primjena algoritama u bankarskom poslovanju - Credit Scoring, Zbornik radova ITeO, Banja Luka, BiH, 2011.
3. Olja Latinović, Ivan Milenković, Dušan Starčević, Velimir Štavljanin, Open Source application in speaker recognition, Zbornik radova Infotech, Vrnjačka Banja, Srbija, 2012.
4. Ivan Milenković, Olja Latinović, Dejan Simić, Identity management in cloud computing, Zbornik radova ITeO, Banja Luka, BiH, 2012.
5. Ivan Milenković, Olja Latinović, Uroš Šošević, Dejan Simić, Primena kriptografije u biometrijskim sistemima, Zbornik radova Infotech, Arandelovac, Srbija, 2013.
6. Olja Latinović, Ivan Milenković, Prototip komunikacionog adaptera za rad sa unimodalnim biometrijskim rešenjem, Zbornik radova ITeO, Banja Luka, BiH, 2013.
7. Olja Latinović, Zoran Marjanović, Speech recognition in Java, SrOUG, Arandelovac, Srbija, 2014.
8. Olja Latinović, Uroš Šošević, Dušan Starčević, Tehnologije otvorenog koda za prepoznavanje govornika, Zbornik radova Infotech, Arandelovac, Srbija, 2015.

Konferencije:

2010 - 2015: Član organizacionog odbora, ITeO, Banja Luka, BiH

Sertifikati:

2013: Oracle Database 11g Performance Tuning Certified Expert

2013: Oracle Database 11g Administrator Certified Professional

2014: Oracle Certified Professional, Java SE 7 Programmer

Spisak položenih predmeta na doktorskim studijama sa ocenama i ESPB bodovima:

Naziv predmeta	Ocena	ESPB
Upravljanje podacima	10 (deset)	10
Razvoj informacionih sistema	10 (deset)	10
Računarske mreže - odabrana poglavlja	10 (deset)	10
Mobilno računarstvo - odabrana poglavlja	10 (deset)	10
Interakcija čoveka i računara - odabrana poglavlja	10 (deset)	10
Interoperabilnost poslovnih sistema i aplikacija	10 (deset)	10
Multimedijalne komunikacije	10 (deset)	10
Metodologija naučno-istraživačkog rada	10 (deset)	10
Sistemi zaštite informacionih sistema	10 (deset)	10

1.3. Ocena podobnosti kandidata za rad na predloženoj temi

Uzimajući u obzir predmete položene na doktorskim studijama, dosadašnje radno iskustvo i ostvarene rezultate iz predložene oblasti istraživanja zaključuje se da je kandidatkinja Olja Latinović kvalifikovana i dovoljno pripremljena da predloženu temu doktorske disertacije samostalno istraži i pruži odgovarajuće naučne i stručne doprinose u razmatranoj oblasti.

2. Predmet i cilj istraživanja

Problem koji se razmatra u predloženoj doktorskoj disertaciji odnosi se na ispitivanje mogućnosti i ograničenja koja poseduju raspoloživa programska rešenja sa otvorenim programskim kodom za prepoznavanje osoba na osnovu glasa, a koja se menjaju u realnim uslovima rada kao rezultat varijacija u interakciji korisnika sa računarskim sistemom. Projektant aplikativnih rešenja za rad na Internetu u realnom vremenu, prilikom korišćenja govora kao biometrijske karakteristike u sistemima menadžmenta identiteta, ova ograničenja treba imati u vidu.

2.1 Predmet istraživanja

Predmet rada ove doktorske disertacije je sistem za prepoznavanje osoba na osnovu glasa realizovan dostupnim rešenjima iz domena otvorenog programskog koda. Istražiće se uticaj pojave razlika u eksploatacionom režimu rada sistema za prepoznavanja govornika, u odnosu režim rada u razvojnoj fazi (obuka sistema), na eksploatacione performanse sistema, a koje nastupaju kao rezultat varijacija u interakciji korisnika sa predmetnim sistemom. Analiziraće se razlike koje mogu nastupiti u više domena interakcije: nepodudarnost u akvizicionim uređajima za snimanje glasa u procesima obuke i korišćenja sistema, nepodudarnost u radnom okruženju u

koje se odvija akvizicija zvučnog zapisa u toku obuke i korišćenja sistema, nepodudarnost u dužini zvučnog zapisa iz procesa obuke i eksploatacionog režima rada, i nepodudarnost u jeziku koji je korisnik koristio tokom obuke, odnosno tokom korišćenja sistema. Za pristupne uređaje u procesu istraživanja koristiće se konvencionalni stoni računar sa priključenim mikrofonom, tablet i smart telefon, imajući u vidu sve češće korišćenje mrežnih aplikacija u režimu mobilnog korisnika. Kao moguća radna okruženja koristiće se profesionalni studio za snimanje zvuka i prostor Laboratorije za multimedijalne komunikacije Fakulteta organizacionih nauka. Govorni zapisi ispitanika biće svrstani u tri skupa po dužini trajanja, i u tri skupa po semantičkoj interpretaciji sadržaja zapisa. Analiziraće se efekat korišćenja dva različita jezika, srpskog i engleskog, u procesu interakcije na performanse sistema za prepoznavanje osoba na osnovu glasa.

2.2 Cilj istraživanja

Osnovni cilj istraživanja je nalaženje odgovora na pitanje koliko uobičajne varijacije u interakciji korisnika sa sistemom za prepoznavanje korisnika na osnovu glasa mogu uticati na degradaciju funkcionalnih performansi takvih sistema, kao gradivnih elemenata menadžmenta identiteta u mrežnim aplikacijama u realnom vremenu. Ispitivanja će se ponoviti na nekoliko dostupnih rešenja sa otvorenim programskim kodom za prepoznavanje osoba na osnovu glasa, a koja implementiraju iste ili različite obradne paradigme za utvrđivanje autentičnosti izvora zvučnog zapisa. Evaluacija i rangiranje dostupnih programskih rešenja će se zasnivati na merenju performansi takvih sistema u različitim uslovima interakcije korisnika sa sistemom, okvirno opisanim u sekciji Predmet rada.

Specifični podciljevi koje treba ostvariti tokom istraživanja su:

- Sistematizacija i prezentacija osnovnih metodoloških postupaka za prepoznavanje osoba na osnovu glasa;
- Analiza funkcionalnih mogućnosti dostupnih softverskih rešenja sa otvorenim programskim kodom za razvoj sistema za prepoznavanje osobe na osnovu glasa;
- Razvoj odgovarajuće metodologije vrednovanja uticaja varijacija u načinu interakcije čoveka i sistema za prepoznavanje govornika na performanse takvih sistema;
- Formiranje radnog okruženja sa akvizicionim uređajima za snimanje zvučnih zapisa: računar sa priključenim mikrofonom, tablet i smart telefon;
- Formiranje više lista tekstualnih zapisa različite dužine (kratki, srednji i duži tekst) na srpskom i engleskom jeziku. Svaka lista tekstova je specifičnog karaktera. Tekstovi iz prve liste treba da budu iz računarskog rečnika, češće korišćenog u komunikaciji posredstvom Interneta. Tekstovi iz druge liste treba da budu iz domena bankarskih transakcija, dok tekstovi iz treće liste obuhvataju svakodnevne aktivnosti ljudi;
- Izbor i obuka dve grupe ispitanika za snimanje zvučnih zapisa. Na osnovu delova prve grupe zvučnih zapisa ispitanika treba formirati u procesu obuke bazu modela ispitanika kao govornika. Na osnovu drugog dela zvučnog zapisa iz prve grupe u eksploatacionom režimu rada treba meriti funkcionalne performanse sistema u različitim uslovima interakcije. Zvučni zapisi iz druge grupe koristiće se za ispitivanje performansi sistema u radu sa uljezima;
- Istovremeno snimanje zvučnih zapisa pročitanih tekstova posredstvom tri akviziciona uređaja u studijskom i laboratorijskom prostoru u više sesija;

- Priprema snimljenih zvučnih zapisa (pretprocesiranje) za dalju obradu imajući u vidu NIST SRE preporuke;
- Na osnovu dostupnih rešenja u otvorenom programskom kodu, formirati više instanci eksperimentalnih laboratorijskih sistema za prepoznavanje osobe na osnovu glasa;
- Specifična obuka više instanci laboratorijskih sistema za prepoznavanje osobe koristeći formirane liste zvučnih zapisa;
- Pripremiti odgovarajuća ispitna scenarija, koja opisuju varijacije u interakciji korisnika i sistema, i na osnovu njih za svaku instancu laboratorijskog eksperimentalnog sistema za prepoznavanje govornika formirati različite ispitne konfiguracije elemenata sistema;
- Na osnovu formiranih ispitnih konfiguracija elemenata sistema na svakoj instanci laboratorijskog sistema za prepoznavanje osoba na osnovu glasa izvršiti eksperimentalno ispitivanje uticaja izloženih varijacija na funkcionalne i nefunkcionalne performanse posmatranih sistema i prikupiti odgovarajuće podatke;
- Na osnovu prikupljenih podataka o ponašanju laboratorijskih sistema za prepoznavanje osoba na osnovu glasa izvršiti odgovarajuću statističku obradu i na pogodan način prikazati rezultate obrade;
- Interpretirati dobijene rezultate u pogledu robustnosti ponaosob za svaku instancu laboratorijskih eksperimentalnih sistema za prepoznavanje osoba na osnovu glasa;
- Evaluirati i rangirati dostupna rešenja za prepoznavanje osoba na osnovu glasa sa otvorenim kodom u smislu pogodnosti korišćenja takvih rešenja u sistemima menadžmenta identiteta.

Okvirna lista relevantnih bibliografskih izvora koji će se koristiti prilikom izrade doktorske disertacije

- [1] A. Anjos, L. El Shafey, R. Wallace, M. Günther, C. Mc-Cool, and S. Marcel, "Bob: a free signal processing and machine learning toolbox for researchers," in ACM International Conference on Multimedia, 2012.
- [2] S. Mokhov, On Design and Implementation of Distributed Modular Audio Recognition Framework: Requirements and Specification Design Document. Department of Computer Science and Software Engineering, Concordia University, 2006, <http://marf.sf.net>.
- [3] P. Kenny and P. Dumouchel, "Disentangling speaker and channel effects in speaker verification," in IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP, pp. 37–40., 2004.
- [4] Steve Young, Gunnar Evermann, Mark Gales, Thomas Hain and others, „HTK Book“ (for HTK 3.4 version), Cambridge University Engineering Department, 2001.
- [5] [D. Matrouf, J.-F. Bonastre, C. Fredouille, A. Larcher, S. Mezaache, M. McLaren, and F. Huenupan. LIA GMM-SVM system description: NIST SRE. In NIST SRE, Montreal (Canada), 2008.
- [6] Von Kriegstein, K., & Giraud, A. L. (2006). Implicit multisensory associations influence voice recognition. *PLoS Biol*, 4(10), e326.
- [7] Kaur, D., Kaur, G., & Singh, D. (2013). Efficient and Robust Multimodal Biometric System for Feature Level Fusion (Speech and Signature). *International Journal of Computer Applications*, 75(5), 33-38.
- [8] Jain, A. K., Hong, L., & Kulkarni, Y. (1999, March). A multimodal biometric system using fingerprint, face and speech. In *Proceedings of 2nd Int'l Conference on Audio-and Video-based Biometric Person Authentication*, Washington DC (pp. 182-187).

- [9] Brady, Kevin, et al. "An evaluation of audio-visual person recognition On the XM2VTS corpus using the lausanne protocols." *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*. Vol. 4. IEEE, 2007.
- [10] Viswanathan, Mahesh, et al. "Information access using speech, speaker and face recognition." *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on*. Vol. 1. IEEE, 2000.
- [11] Roy, D., Shukla, A., "Speaker recognition using multimodal biometric system," *Oriental COCOSDA held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE), 2013 International Conference* , vol., no., pp.1,7, 25-27 Nov. 2013
- [12] Çetingül, Hasan Ertan, et al. "Multimodal speaker/speech recognition using lip motion, lip texture and audio." *Signal processing* 86.12 (2006): 3549-3558.
- [13] R.Klevansand R.Rodman, "Voice Recognition, Artech House, Boston, London 1997.
- [14] GIN-DER WU AND YING LEI " A Register Array based Low power FFT Processor for speech recognition" Department of Electrical engineering national Chi Nan university Puli ,545 Taiwan, 2008.
- [15] Sadoki Furuki, Tomohisa Ichiba et. al, Cluster-based Modeling for Ubiquitous Speech Recognition, Department of Computer Science Tokyo Institute of Technology Interspeech 2005.
- [16] Singer, E., Torres-Carrasquillo, P. A., Gleason, T. P., Campbell, W. M., & Reynolds, D. A. (2003, September). Acoustic, phonetic, and discriminative approaches to automatic language identification. In *INTERSPEECH*.
- [17] Le, V. B., Besacier, L., & Schultz, T. Acousticphonetic unit similarities for context dependant acoustic model portability Carnegie Mellon University.*Pittsburgh, PA, USA*.
- [18] Morales, Nicolás, John HL Hansen, and Doroteo Torre Toledano. "MFCC Compensation for Improved Recognition of Filtered and Band-Limited Speech." *ICASSP (1)*. 2005.
- [19] Razak, Zaidi, et al. "Quranic verse recitation feature extraction using mel-frequency cepstral coefficient (MFCC)." *4th International Colloquium on Signal Processing and its Applications (CSPA 2008)*. 2008.
- [20] W. M. Campbell_, D. E. Sturim W. Shen D. A. Reynolds_, J. Navr'atily "The MIT-LL/IBM Speaker recognition System using High performance reduced Complexity recognition" MIT Lincoln Laboratory IBM 2006.
- [21] Karam, Zahi N., and William M. Campbell. "A new kernel for SVM MLLR based speaker recognition." *INTERSPEECH*. 2007.
- [22] Asghar Taheri, Mohammad Reza Trihi et. al, Fuzzy Hidden Markov Models for speech recognition on based FEM Algorithm, Transaction on engineering Computing and Technology V4 February 2005,IISN,1305-5313
- [23] Alexander, Anil, et al. "Speaker recognition with Phonetic and Automatic Features using VOCALISE software." *22th Annual Conference of the International Association for Forensic Phonetics and Acoustics, Tampa*. 2013.
- [24] Greenberg, Craig S., et al. "The NIST 2014 speaker recognition i-vector machine learning challenge." *Odyssey: The Speaker and Language Recognition Workshop*. 2014.
- [25] Charton, Eric, et al. "Mistral: open source biometric platform." *Proceedings of the 2010 ACM Symposium on Applied Computing*. ACM, 2010.

- [26] Dehak, N.; Kenny, P.; Dehak, R.; Dumouchel, P.; Ouellet, P., "Front-End Factor Analysis for Speaker Verification," *Audio, Speech, and Language Processing, IEEE Transactions on* , vol.19, no.4, pp.788,798, May 2011
- [27] Forsberg, Markus. "Why is speech recognition difficult." Chalmers University of Technology (2003).
- [28] F. Jelinek, "Continuous Speech Recognition by Statistical Methods." *IEEE Proceedings* 64:4: 532-556, 1976.
- [29] Young, S., A Review of Large-Vocabulary Continuous Speech Recognition, *IEEE Signal Processing Magazine*, pp. 45-57, Sep. 1996
- [30] Garcia-Romero, D.; Xinhui Zhou; Espy-Wilson, Carol Y., "Multicondition training of Gaussian PLDA models in i-vector space for noise and reverberation robust speaker recognition," *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on* , vol., no., pp.4257,4260, 25-30 March 2012
- [31] A.K. Jain, A. Ross, S. Prabhakar, "An introduction to biometric recognition", *IEEE Trans. Circuits Systems Video Technol.* 14(1) 4–20, 2004.
- [32] Sampath, Ashwin, Holtzman, J.M., "Access control of data in integrated voice/data CDMA systems: benefits and tradeoffs," *Selected Areas in Communications, IEEE Journal on* , vol.15, no.8, pp.1511,1526, 1997.
- [33] Feix, Wolfgang H., and Hartwig E. Ruell. "Personal access control system using speech and face recognition." U.S. Patent No. 4,449,189., 1984.
- [34] Dowling, Kevin J., and George G. Mueller. "Lighting control using speech recognition." U.S. Patent No. 7,031,920. 18 Apr. 2006.
- [35] Scott, Brian L., and Clint L. Miller. "Using speech recognition to access the internet, including access via a telephone." U.S. Patent No. 6,101,473. 8 Aug. 2000.
- [36] Olja Latinović, Ivan Milenković, Dušan Starčević, Velimir Štavljanić, *Primena rešenja tehnologije otvorenog koda u prepoznavanju osobe zasnovanom na glasovnom zapisu*, Infotech, Srbija, 2012
- [37] T. Kinnunen and P. Rajan, "A practical, self-adaptive voice activity detector for speaker verification with noisy telephone and microphone data," in *IEEE ICASSP*, 2013.
- [38] Kinnunen, Tomi, and Haizhou Li. "An overview of text-independent speaker recognition: From features to supervectors." *Speech communication* 52.1 (2010): 12-40.
- [39] Larcher, Anthony, et al. "ALIZE 3.0-open source toolkit for state-of-the-art speaker recognition." *INTERSPEECH*. 2013.
- [40] Khoury, Elie, Laurent El Shafey, and Sébastien Marcel. "Spear: An open source toolbox for speaker recognition based on Bob." *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014.
- [41] Buzo, Andi, et al. "An Automatic Speech Recognition solution with speaker identification support." *Communications (COMM), 2014 10th International Conference on*. IEEE, 2014.
- [42] Thakur, Ajay Kumar, et al. "Case Study Of GIPSY and MARF." *arXiv preprint arXiv:1407.3347*, 2014.
- [43] Lelong, Amélie, and Gérard Bailly. "Characterizing phonetic convergence with speaker recognition techniques." *The Listening Talker Workshop (LISTA 2012)*. 2012.
- [44] Liu, Qiuge. "SYSTEM AND METHOD FOR LOAD BALANCING IN A SPEECH RECOGNITION SYSTEM." U.S. Patent No. 20,140,337,022., 2014.
- [45] Candini, Michela, et al. "Who is speaking? Implicit and explicit self and other voice recognition." *Brain and cognition* 92: 112-117., 2014.

- [46] Navya, M., S. Mohammed Rafi, and K. Niranjan Reddy. "Android Based Children Tracking System Using Voice Recognition.", 2015.
- [47] Rosenberg, A. E., Lee, C. H., & Gokcen, S. (1991, April). Connected word talker verification using whole word hidden Markov models. In *Acoustics, Speech, and Signal Processing, ICASSP-91., 1991 International Conference on* (pp. 381-384). IEEE, 1991.
- [48] Stan Z. Li – "Encyclopedia of Biometrics", Springer, 2009.
- [49] Peter Gregory, Michael A. Simon – "Biometrics for Dummies", Wiley Publishing, Inc., 2008.
- [50] Vogt R., Sridharan S. – „Explicit modeling of session variability for speaker verification“, Computer Speech and Language 22, 1, 2008.
- [51] Wolf J. – „Efficient acoustic parameters for speaker recognition“, Journal of the Acoustic Society of America, 1972.
- [52] B. S. Atal - "Automatic Recognition of Speakers from their Voices", Proceedings of the IEEE, 1976.
- [53] Addy, Kenneth L., and Ronald K. Rothman. "System and method of speaker recognition." U.S. Patent No. 8,971,854., 2015..
- [54] Rao, K. Sreenivasa, and Sourjya Sarkar. "Robust Speaker Verification: A Review." Robust Speaker Recognition in Noisy Environments. Springer International Publishing, 13-27., 2014.
- [55] Ma, Bin, Helen M. Meng, and Man-Wai Mak. "Effects of device mismatch, language mismatch and environmental mismatch on speaker verification." Acoustics, speech and signal processing, 2007. ICASSP 2007. IEEE International Conference on. Vol. 4. IEEE, 2007.
- [56] Bimbot, Frédéric, et al. "A tutorial on text-independent speaker verification." EURASIP journal on applied signal processing 2004: 430-451, 2004.
- [57] Reynolds, Douglas A., Thomas F. Quatieri, and Robert B. Dunn. "Speaker verification using adapted Gaussian mixture models." Digital signal processing 10.1 :19-41., 2000.
- [58] Biometrics Metrics Report v3. 0, Prepared for U. S. Military Academy (USMA) -West Point, p. 5, 2012. <http://www.usma.edu/ietd/docs/BiometricsMetricsReport.pdf>
- [59] Jain, A. K., Bolle, R., & Pankanti, S. Biometrics: Personal Identification. Dordrecht: Kluwer Academic Publishers, 1999.
- [60] Li, S. Z. (2009). Encyclopedia of Biometrics (1st ed.). New York, NY: Springer Publishing Company, Incorporated.
- [61] Reid, Paul. Biometrics for network security. Prentice Hall Professional, 2004.
- [62] Velibor Ilić, Neuronske mreže, 1998.

3. Polazne hipoteze istraživanja

Glavna hipoteza koja će biti testirana u radu glasi:

Eksploatacione performanse dostupnih programskih rešenja za prepoznavanje osoba na osnovu glasa su veoma osetljive na varijacije u načinu interakcije korisnika sa sistemom.

Na osnovu definisanog predmeta istraživanja može se izdvojiti i nekoliko posebnih hipoteza:

Pomoćna hipoteza 1: Degradacija funkcionalnih performansi sistema za prepoznavanje osobe na osnovu glasa sa rešenjem u otvorenom kodu je posledica nepodudarnosti pristupnih uređaja kao delova korisničkog interfejsa u procesima obuke i eksploatacije sistema.

Pomoćna hipoteza 2: Degradacija funkcionalnih performansi sistema za prepoznavanje osoba je posledica nepodudarnosti radnog okruženja u kojem se odvija komunikacija korisnika sa sistemom tokom obuke i tokom eksploatacije sistema.

Pomoćna hipoteza 3: Degradacija funkcionalnih performansi sistema za prepoznavanje osoba je posledica nepodudarnosti jezika na kojem korisnik komunicira sa sistemom tokom obuke i tokom eksploatacije sistema.

Pomoćna hipoteza 4: Degradacija funkcionalnih performansi sistema za prepoznavanje osoba je posledica nepodudarnosti dužine trajanja komunikacije korisnika sa sistemom tokom obuke, odnosno tokom eksploatacije sistema.

Pomoćna hipoteza 5: Degradacija funkcionalnih performansi sistema za prepoznavanje osoba je posledica semantičke nepodudarnosti zvučnih zapisa korisnika tokom obuke i tokom eksploatacije sistema.

4. Naučne metode istraživanja

Tokom navedenog istraživanja koristiće se više naučnih metoda:

- Prikupljanje i sređivanje podataka o dostupnim relevantnim rešenjima na osnovu dostupne literature, pretraživanje Interneta, razmenom informacija i neposrednim kontaktima sa relevantnim subjektima u svetu;
- Kritička analiza korišćenih metodologija razvoja, alata i tehnologija u domenu prepoznavanja govornika radi uočavanje dobrih i loših osobina;
- Sinteza, apstrakcija, konkretizacija, generalizacija, specijalizacija i klasifikacija, kao i ostale metode naučnog objašnjenja, biće primenjene na osnovu indukcije i dedukcije;
- Deduktivno-logička metoda pomoću koje će se na temelju spoznaje utvrđene praksom, utemeljiti vlastiti teorijski stavovi;
- Postavljanje sopstvene metodologije za ostvarivanje postavljenih ciljeva istraživanja sintezom poznatih i proverenih elemenata rešenja, ili njihovom modifikacijom, sa originalnim elementima rešenja;
- Istraživanja korišćenjem statističke metode prikupljanja uzoraka;
- Deskriptivna analiza i statističke metode za obradu podataka i prezentaciju rezultata;
- Eksperimentalna provera predložene metodologije istraživanja realizacijom laboratorijske verzije rešenja.

5. Očekivani naučni doprinos

Očekivani rezultati istraživačkog procesa u predmetnoj oblasti treba da sadrže više značajnih naučnih doprinosa, i to:

- Celovit kritički pregled istraživačke oblasti prepoznavanja osoba na osnovu glasa i sistematizacija postojećih pristupa,
- Razvoj odgovarajuće metodologije za vrednovanje uticaja varijacija u procesu interakcije korisnika i računarskog sistema na performanse sistema za prepoznavanje govornika,
- Formiranje originalne biometrijske baze podataka sa zvučnim zapisima u skladu sa zahtevima postavljene metodologije,
- Formiranje eksperimentalnog laboratorijskog računarskog sistema sa više instanci programskih rešenja za prepoznavanje osobe na osnovu glasa, usklađenog sa zahtevima metodologije vrednovanja uticaja varijacija u načinu interakcije korisnika i takvih sistema,
- Semantička interpretacija podataka prikupljenih na eksperimentalnom laboratorijskom sistemu, u skladu sa postavljenom metodologijom vrednovanja uticaja varijacija u načinu interakcije, i potvrda ili odbacivanje postavljenih hipoteza istraživanja.

Stručni doprinos se može očekivati postavljenjem odgovarajućih smernica projektantima sistema menadžmenta identitetom, zasnovanog na prepoznavanju korisnika na osnovu glasa, u vezi korišćenja raspoloživih programskih rešenja sa otvorenim programskim kodom.

Društveni doprinos se ogleda u predočavanju društvenoj zajednici mogućnosti primene dela biometrijskih tehnologija u izgradnji bezbedonosnih mehanizama za upravljanje pristupom zaštićenim računarskim resursima.

6. Plan istraživanja i struktura rada

Istraživanje će se sprovesti pregledom relevantne literature, analizom rezultata na realizaciji korporativnog sistema obrazovanja u konkretnom poslovnom sistemu i profesionalnoj organizaciji, kao i predlogom njihovog daljeg unapređenja. Posebna pažnja u istraživanju posvetiće se prikupljanju stavova zaposlenih, kao i njihovoj statističkoj obradi i analizi.

Plan istraživanja doktorske disertacije prikazan je u sledećoj tabeli:

[illegible]

Okvirni sadržaj predložene doktorske disertacije

1. Uvod
2. Menadžment identiteta i biometrijski sistemi identifikacije
 - 2.1. Menadžment identiteta
 - 2.2. Biometrijski sistemi identifikacije
 - 2.3. Sistemi identifikacije osoba na osnovu glasa
3. Pregled pristupa sistemima za prepoznavanje govornika
 - 3.1. Tekstualno zavisni i tekstualno nezavisni sistemi prepoznavanja govornika
 - 3.2. Opšta arhitektura sistema za prepoznavanje govornika
 - 3.3. Izvođenje karakteristika glasa
 - 3.4. Modeli govornika
 - 3.5. Pristupi problemu robustnosti
 - 3.6. Kritički osvrt na postojeće pristupe
4. Pregled programskih rešenja za prepoznavanje govornika sa otvorenim programskim kodom
 - 4.1. SPEAR
 - 4.2. MARF
 - 4.3. Alize
 - 4.4. HTK
 - 4.5. Mistral
 - 4.6. Sumarni prikaz funkcionalnosti programskih rešenja
5. Metodologija vrednovanja uticaja varijacija u načinu interakcije korisnika sa sistemom na performanse sistema za prepoznavanje govornika
6. Formiranje baze zvučnih zapisa
 - 6.1. Radno okruženje za formiranje baze zvučnih zapisa
 - 6.2. Formiranje tekstualnih lista za bazu zvučnih zapisa
 - 6.3. Izbor i obuka ispitanika
 - 6.4. Snimanje zvučnih zapisa
 - 6.5. Pretprocesiranje zvučnih zapisa
 - 6.6. Organizacija i skladištenje pretprocesiranih zvučnih zapisa
7. Empirijska provera predloženog postupka vrednovanja uticaja varijacija u interakciji korisnika i sistema
 - 7.1. Formiranje eksperimentalnog laboratorijskog sistema sa više instanci sistema za prepoznavanje govornika
 - 7.2. Obuka instanci sistema za prepoznavanje govornika
 - 7.3. Izrada scenarija mogućih varijacija u interakciji u cilju ispitivanja eksploatacionih performansi sistema za prepoznavanje govornika
 - 7.4. Programiranje izvršnih konfiguracija na osnovu scenarija
 - 7.5. Eksperimentalno ispitivanje uticaja varijacija na performanse sistema prema postavljenom scenariju
 - 7.6. Rezultati statističke obrade prikupljenih podataka
8. Evaluacija dobijenih rezultata
 - 8.1. Interpretacija rezultata dobijenih u okviru ispitivanih programskih sistema za prepoznavanje govornika sa otvorenim programskim kodom
 - 8.2. Rangiranje ispitivanih programskih sistema za prepoznavanje govornika sa otvorenim programskim kodom

- 9. Zaključak
 - 9.1. Ostvareni naučni i stručni doprinosi
 - 9.2. Mogućnosti primene
 - 9.3. Mogući dalji pravci istraživanja
- 10. Literatura

7. **Zaključak i predlog**

Na osnovu izloženog može se zaključiti da kandidatkinja Olja Latinović ispunjava sve uslove predviđene Zakonom o visokom obrazovanju za odobrenje izrade doktorske disertacije pod naslovom “ Uticaj varijacija u interakciji korisnika na performanse sistema za prepoznavanje govornika“.

Tema pripada užoj naučnoj oblasti Informacione tehnologije. Ostvareni naučni stručni rezultati mogu imati širu primenu u širenju i unapređenju menadžmenta identiteta u informacionim sistemima.

Na osnovu svega navedenog, komisija predlaže Nastavno-naučnom veću da prihvati predloženu temu i odobri izradu prijavljene doktorske disertacije. Za mentora doktorske disertacije predlaže se dr Dušan Starčević, redovni profesor Fakulteta organizacionih nauka, Univerziteta u Beogradu.

U Beogradu, 12.10.2015.

ČLANOVI KOMISIJE

dr Dušan Starčević, red. prof. Fakulteta organizacionih
nauka Univerziteta u Beogradu

dr Miroslav Minović, docent Fakulteta organizacionih
nauka Univerziteta u Beogradu

dr Zoran Jovanović, red. prof. Elektrotehničkog fakulteta
Univerziteta u Beogradu