

Образац за пријаву техничког решења¹

Назив техничког решења	Напредна детекција говорних сегмената на мобилном телефону применом микрофонског низа и декодера
Аутори техничког решења	Бранислав Поповић, Едвин Пакоци, Дарко Пекар, Владо Делић
Категорија техничког решења	Ново техничко решење (метода) примењено на националном нивоу (M82)

За кога је рађено техничко решење и у оквиру ког пројекта МПНТР:

Техничко решење је реализовано у склопу пројекта технолошког развоја “Развој дијалošких система за српски и друге јужнословенске језике” (ТР32035, 2011-2016) на Факултету техничких наука и у предузећу АлфаНум у Новом Саду.

Ко користи техничко решење:

Предузеће АлфаНум у Новом Саду, као партиципant на пројекту ТР32035, као и предузеће Ахон - интелигентни системи у Новом Саду.

Година када је техничко решење урађено:

Развој методе за детекцију говорних сегмената у условима повећаног нивоа буке почео је у другој половини 2015. године, а завршен је у другој половини 2016. године. Говорне технологије примењене у склопу техничког решења развијене су у периоду од 2005. до 2016. године, у оквиру технолошких пројеката МПНТР. За потребе обуке и тестирања препознавача говора на великим речницима коришћене су говорне базе прикупљене у периоду од 2014. до 2016. године.

Ко је прихватио-примењује техничко решење:

Ово техничко решење иницијално је реализовано за потребе развоја говорних технологија у предузећу АлфаНум, а користи се и у развојном тиму на Факултету

¹ У складу са одредбама *Правилника о поступку, начину вредновања и квантитавном исказивању научноистраживачких резултата истраживача*, донетог од стране министра просвете, науке и технолошког развоја на основу члана 70. ст. 7. и 8. и члана 86. став 2. Закона о научноистраживачкој делатности („Службени гласник РС”, бр. 110/05, 50/06-исправка, 18/10 и 112/15).

техничких наука у Новом Саду. Метода је након тога примењена у предузећу Ахон - интелигентни системи, у оквиру апликације за управљање мобилним телефоном путем гласа на српском језику.

Како су резултати верификовани (од стране ког тела):

- 1) Техничко решење је реализовано у Лабораторији за акустику и говорне технологије на Факултету техничких наука у Новом Саду. Имплементирано је и испитано на развојним системима у предузећу АлфаНум. Користи се у оквиру апликације за управљање мобилним телефоном путем гласа на српском језику у предузећу Ахон - интелигентни системи у Новом Саду.
- 2) Техничко решење је ново техничко решење (метода), примењено на националном нивоу, засновано на више техничких и развојних решења која су пријављивана као резултати на технолошким и иновационим пројектима МПНТР у периоду од 2005. до 2016. године. Решење је реализовано у оквиру пројекта TP32035 („Развој дијалогских система за српски и друге јужнословенске језике“). Опис предложене методе и резултати примене приказани су у раду
 - Бранислав Поповић, Едвин Пакоци, Дарко Пекар, „Advanced Voice Activity Detection on Mobile Phones by Using Microphone Array and Phoneme-Specific Gaussian Mixture Models“, *IEEE 14th International Symposium on Intelligent Systems and Informatics (SISY 2016)*, Суботица, Србија, 29-31. август 2016, pp. 45-48, ISBN: 978-1-5090-2866-5.

Поједини елементи техничког решења презентовани су на неколико међународних научних скупова, на пример:

- Бранислав Поповић, Едвин Пакоци, Никша Јаковљевић, Горан Кочиш, Дарко Пекар, „Voice Assistant Application for the Serbian Language“, *23rd Telecom. Forum, TELFOR-2015*, Београд, Србија, 24-26. новембар 2015, pp. 858-861, DOI: 10.1109/TELFOR.2015.7377600, ISBN: 978-1-5090-0054-8.
 - Бранислав Поповић, Едвин Пакоци, Никша Јаковљевић, Дарко Пекар, Стеван Острогонац, „A Robust Voice Activity Detection in Voice Assistant Application“, *Book of Abstracts, 3rd International Acoustic and Audio Engineering Conference, TAKTONS 2015*, Нови Сад, Србија, 18-21. новембар 2015, pp. 38-39, ISBN: 978-86-7892-758-4.
 - Бранислав Поповић, Стеван Острогонац, Едвин Пакоци, Никша Јаковљевић, Владо Делић, „Deep Neural Network Based Continuous Speech Recognition for Serbian Using the Kaldi Toolkit“, *17th SPECOM, Speech and Computer, Атина, Грчка, 20-24. септембар 2015, Lecture Notes in Computer Science, Springer, vol. 9319, pp. 186-192, DOI: 10.1007/978-3-319-23132-7_23, Series ISSN: 0302-9743, Print ISBN: 978-3-319-23131-0, Online ISBN: 978-3-319-23132-7.*
- 3) Дато је писано мишљење два рецензента - експерта из области техничког решења:
 - Др Јернеја Жганец Грос, Алпинеон д.о.о. (Љубљана, Словенија)
 - Др Бранислав Геразов, Факултет за електротехнику и информационе технологије (Скопље, Македонија)

- 4) Наставно-научно веће Департмана за енергетику, електронику и телекомуникације и Факултета техничких наука, на основу мишљења рецензената и приложених доказа, издало је Уверење о признавању техничког решења које потврђује да оно испуњава све услове да буде признато као ново техничко решење (метода) примењено на националном нивоу (M82), у складу са Правилником Министарства.

На који начин се користи (кратак опис):

Техничко решење представља нову методу за напредну детекцију говорних сегмената на мобилном телефону применом микрофонског низа и декодера. Примењује се у оквиру апликације за управљање мобилним телефоном путем гласа на српском језику. Поменута апликација је развијена за мобилне платформе засноване на Андроид оперативном систему. Техничко решење је реализовано у виду модула за детекцију говорних сегмената (*voice activity detection* – VAD) и пружа знатно увећану робустност система на присуство позадинског шума, што је посебно важно за примену система за препознавање говора у покрету (нпр. у бучним канцеларијама, при управљању аутомобилом, на улици и слично). Сврха овог техничког решења је да омогући брзу и поуздану детекцију говорне активности од интереса, посебно у условима повишеног нивоа буке. Применом микрофонског низа у оквиру предложеног модула за потискивање шума, као и декодера са ограниченом граматиком за детекцију говора, у комбинацији са енергијом и акустичким тежинама, остварене су високе перформансе система, чак и у случајевима када је присутан висок ниво позадинског шума. Примена техничког решења истовремено доводи и до значајног побољшања брзине одзива, имајући у виду да се само они делови сигнала у којима је детектована говорна активност шаљу на даље препознавање, при чему се користе акустички модели обучени за препознавање говора у условима повећаног нивоа шума.

Опис техничког решења: Напредна детекција говорних сегмената на мобилном телефону применом микрофонског низа и декодера

Техничко решење омогућава напредну детекцију говорних сегмената на мобилном телефону и представља важан део постојећег система за управљање мобилним телефоном путем гласа на српском језику, реализованог у виду апликације за мобилне уређаје засноване на Андроид оперативном систему. Решење је остварено применом софтверских решења и напредних говорних технологија за српски и друге јужнословенске језике, развијеним коришћењем научних метода на пројектима МПНТР. Развој модула за детекцију говорних сегмената на мобилном телефону пратио је већи број научних и стручних радова, као и развојних и техничких решења, који су пријављивани као резултати на пројектима МПНТР у периоду 2005-2016. године. По својој комплексности, архитектури, дизајну и функционалностима које пружа, ово

техничко решење је јединствено на српском говорном подручју. Примена техничког решења успешно је демонстрирана на неколико научних и стручних скупова у земљи и иностранству.

Област на коју се техничко решење односи:

Техничко решење припада области Електроника, телекомуникације и информационе технологије, и представља спој различитих области, попут телекомуникација и обраде сигнала, дигиталне обраде звука, вештачке интелигенције, препознавања говора и шума, и других.

Проблем који се техничким решењем решава:

Детекција говорних сегмената спада међу најважније задатке у системима за аутоматско препознавање говора. Робустност и квалитет детекције говорних сегмената значајно утиче на перформансе система за препознавање говора, како у погледу брзине, тако и у погледу тачности. Ово је посебно важно када је реч о уређајима ограничене процесорске моћи, као што су мобилни телефони, таблети, „паметни“ сатови и њима слични. У случају повишеног нивоа буке, детекција делова сигнала од интереса омогућава систему да одреди делове улазног сигнала који садрже корисничке команде, захтеве и упите, који се након тога шаљу преосталим деловима система на даљу обраду. На тај начин, врши се раздвајање делова сигнала од интереса, при чему се већи део сигнала, односно део који садржи само шум, обрађује у оквиру модула који захтевају мању процесорску снагу, а само мањи део у процесорски знатно захтевнијим модулима.

Сврха овог техничког решења је да омогући контролу мобилних уређаја чији је рад заснован на Андроид платформи путем гласа на српском језику, у условима повишеног нивоа буке. Решење представља спој напредних софтверских решења и говорних технологија за српски језик, при чему је обезбеђен једноставан и функционалан интерфејс према базном C++ коду. Применом овог техничког решења остварује се поуздан рад уз високе перформансе у условима повишеног нивоа шума, независно од карактеристика гласа говорника и спољашњих фактора. Истовремено, препознавање делова сигнала од интереса омогућава знатну уштеду у погледу хардверских захтева и рад апликације у реалном времену, као и повећање тачности препознавања, јер се на даљу обраду шаљу само делови сигнала у којима је детектован говор, уз одговарајућу контролу квалитета.

Стање решености тог проблема у свету:

У литератури се данас могу пронаћи различите методе за детекцију говорних сегмената у аудио сигналу. Већина предложених метода заснована је на примени енергетских прагова (2010, K. Woo et al.), спектралних (2013, S. Mousazadeh) или

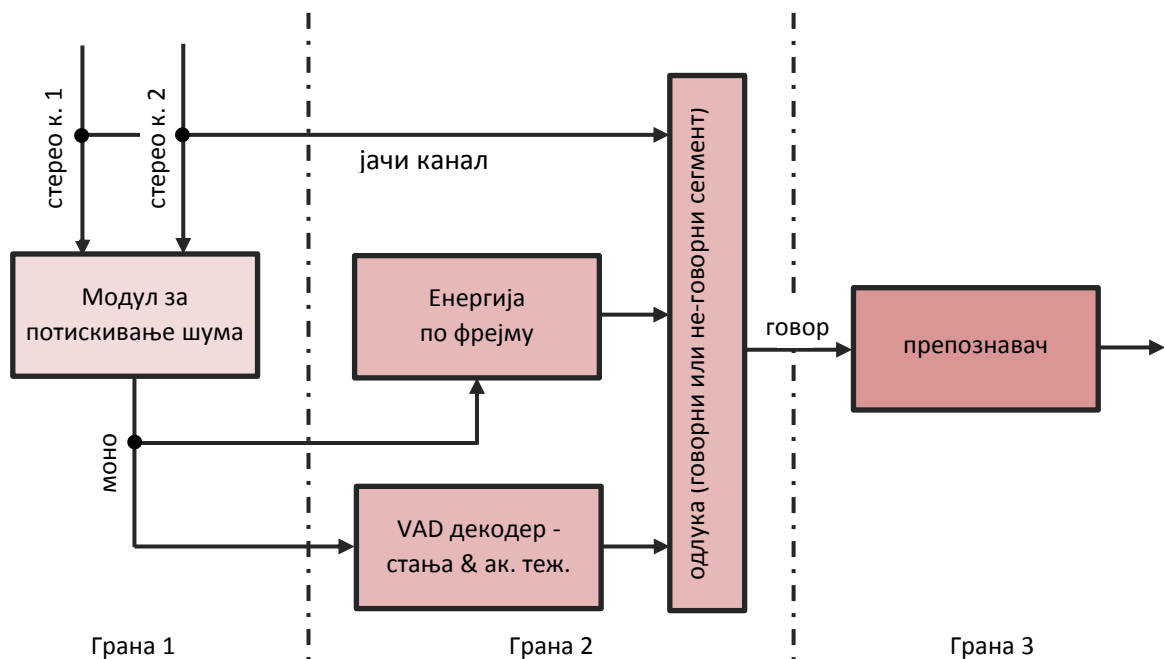
кепстралних обележја (2011, H. Wang et al.), статистика вишег реда (2005, K. Li et al.), временских обележја (2014, B. Liu et al.), аудио-визуелних обележја (2015, S. Montazzolli et al.), модела Гаусових смеша (2011, D. Ying et al.), скривених Марковљевих модела (2012, H. Veisi & H. Sameti), неуронских мрежа (2016, Y. Obuchi) и слично. Методе засноване на примени вектора обележја, нарочито када је реч о енергији, обезбеђују задовољавајуће перформансе у случају чистог сигнала или стационарног шума. Међутим, перформансе таквог система рапидно опадају у случају повећања нивоа шума. Спектрална обележја и статистички модели пружају нешто боље перформансе и генерално се може рећи да су робустнији и ефикаснији (2013, S. O. Sadjadi & J. H. L. Hansen). Модул за детекцију говорних сегмената описан у овом техничком решењу комбинује информације о највероватнијем стању (фонему) за дати одсечак сигнала (фрејм) са акустичком тежином и енергијом, чиме се добијају високе перформансе система за препознавање говора, чак и у присуству значајног нивоа позадинског шума. Применом модула за потискивање шума (2001, M. Brandstein & D. Ward), снижава се ниво позадинског шума у сигналу, чиме се остварује повећана тачност система.

Објашњење суштине техничког решења и детаљан опис са карактеристикама, укључујући и пратеће илустрације и техничке цртеже (техничке карактеристике):

Техничко решење представља нову методу за детекцију говорних сегмената, реализовану у виду модула приказаног на слици 1. Функционалности овог модула су подељене и извршавање тече у 3 паралелне гране (енгл. *thread*), чиме се остварује значајно убрзање комплетног система. Комуникација између ових грана одвија се путем редова чекања (енгл. *buffer*) и одговарајуће сигнализације (семафора). У оквиру прве гране, врши се учитавање стерео сигнала, снимљеног путем микрофонског низа на мобилном телефону, или дуплираног моно канала, у случају да такав низ не постоји. Ова грана обједињује функције чија је сврха потискивање шума, раздвајајући их од осталих делова система за препознавање говора. Излаз из ове гране је моно аудио сигнал, који се путем редова чекања, заједно са оригиналним сигналом из јачег аудио канала, преноси другој грани. Први сигнал се користи за детекцију говорних сегмената, док се релевантни делови другог сигнала (делови сигнала који садрже говор) преносе до препознавача, чија је функционалност заокружена у оквиру треће гране.

У оквиру друге гране, смештен је декодер за издвајање говорних сегмената улазног сигнала. Обука декодера извршена је у 3 фазе. Прва фаза је фаза са применом монофона, при чему се у оквиру излазног модела користи 500 гаусијана. У оквиру друге и треће фазе, користе се модели трифона, са 500 стања и 3500 гаусијана на излазу из друге, односно 900 стања и 7000 гаусијана на излазу из треће фазе. Овај декодер користи граматику у којој су фонеме груписани по класама, при чему се свака класа користи као оригинални фонем (дакле, уместо појединачних фонема, користимо ознаке за тишину - шум, вокале, латералне и назалне фонеме, пловиве, африкате,

фрикативе и полувокал P). Декодер враћа само фонеме (ознаке) од интереса, а то су вокал, полувокал и тишина - шум. У случају да се вокал или полувокал налазе у листи могућих стања за дати одсечак сигнала, стање са најнижом акустичком тежином (ниже вредности су боље) биће проглашено за препознато стање, при чему ће тежина тог стања бити увећана за разлику између стварне вредности и вредности тежине највероватнијег стања у датом моменту. Применом овакве граматике остварене су значајне уштеде по питању времена извршавања, при чему су перформансе модула за детекцију говора остале на задовољавајућем нивоу.

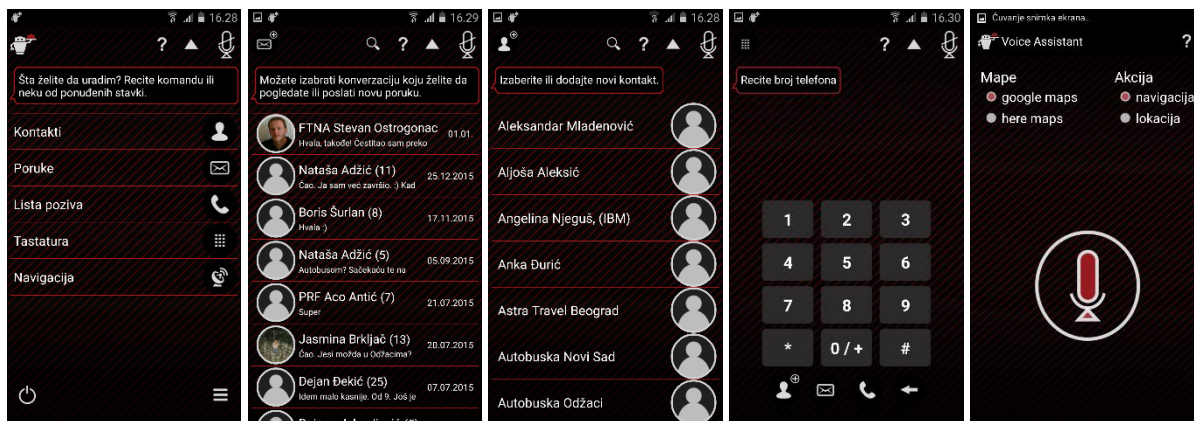


Слика 1. Приказ модула за детекцију говорних сегмената (гране 1 и 2)

Поред декодера, за препознавање говорних сегмената користе се и енергија и акустичке тежине. Оптималне вредности најнижег прага за енергију и највишег прага за акустичку тежину утврђене су експериментално. Применом енергије вршимо раздвајање довољно гласних сигнала од позадинског шума, на пример, корисничких упита од позадинског шума, попут музике или говора у позадини. На тај начин повећава се робустност система јер је модул за потискивање позадинског шума непоуздан у ситуацији када се користе моно сигнали, или нам нежељени говор долази из истог смера као одговарајућа команда. Акустичке тежине служе да раздвоје говор од не-говора, дајући вероватноће препознатих стања, чак и у случајевима када је ниво позадинске буке висок. За детекцију говора потребан је одговарајући број sukcesивних говорних фрејмова (обично 3 или 4), као и за детекцију тишине, при чему је захтевани број фрејмова знатно већи (око 100). Да би се избегле ситуације са краткотрајним бучним догађајима у позадини, дозвољена су до 2 узастопна фрејма говора у оквиру стотинак фрејмова тишине. Притом се за детекцију тишине користе релативни прагови,

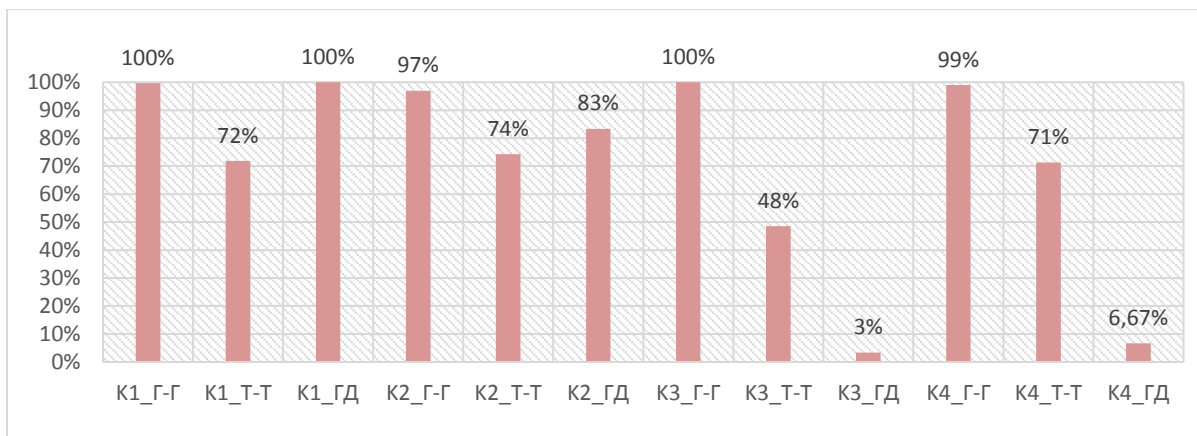
Документација за техничко решење: Напредна детекција говорних сегмената на мобилном телефону применом микрофонског низа и декодера

тј. акустичка тежина мора бити већа од минималне просечне тежине говорног фрејма одређене за дати фрејм и претходна 2 фрејма, за комплетну историју говорног дела сигнала, плус вредност прага за акустичку тежину, док енергија мора бити нижа од максималне просечне енергије одређене за дати фрејм и претходна 2 фрејма, за комплетну историју говорног дела сигнала, минус вредност прага за енергију.



Слика 2. Изглед појединих прозора апликације у оквиру које је интегрисано техничко решење

На слици 2 приказан је изглед појединих прозора апликације у оквиру које је интегрисан модул за детекцију говорних сигнала. Препознавач користи пуну листу фонема и захтева знатно већу процесорску снагу, па се захваљујући примени предложеног модула на препознавање шаљу само делови сигнала у којима је детектован говор. У оба случаја користе се уобичајена обележја - енергија и 14 Мел-фреквенцијских кепстралних коефицијената, као и њихови први и други изводи.



Слика 3. Резултати примене модула за детекцију говорних сегмената

На слици 3 приказани су експериментални резултати, остварени применом предложеног модула у оквиру горе описаног система, одређени за базу од укупно 120 аудио сигнала, подељених у 4 категорије. Прва и друга категорија садрже сигнале од

интереса (команде и упите), при чему прва садржи чисте сигнале (без позадинског шума), док друга садржи знатну количину позадинског шума. Трећа и четврта категорија садрже сигнале које не треба детектовати. Трећа категорија обухвата чист (позадински) говор, док четврта уз позадински говор обухвата и знатну количину шума. Сигнали су снимљени применом апликације посебно дизајниране за ове намене, при чему је размак између уста говорника и мобилног телефона ограничен на 10 до 20 cm за прве две категорије сигнала, односно на преко 50 cm за последње две. K1 до K4 су ознаке категорија, Г-Г представља проценат говорних фрејмова који су детектовани као говорни фрејмови, у односу на укупан број фрејмова који садрже говор. Т-Т представља број фрејмова тишине (не-говора) који су детектовани као фрејмови тишине, у односу на укупан број фрејмова тишине. Обе вредности су израчунате само за оне сигнале унутар којих је детектован говорни сегмент. Проценат таквих аудио сигнала означен је са ГД.

Показује се да предложени скуп обележја даје најбоље резултате у поређењу са системима који поменута обележја користе појединачно. Док енергија обезбеђује разликовање сигнала у прве две у односу на друге две категорије, али уз недовољну тачност, акустичке тежине показују већу тачност на чистом делу базе, али нису довољно дистинктивне. У комбинацији са декодером за препознавање говорних сегмената постижу се оптимални резултати.

Како је реализовано и где се примењује, односно које су могућности примене (техничке могућности):

Техничко решење представља ново техничко решење, примењено у виду модула интегрисаног у оквиру апликације за управљање мобилним телефоном путем гласа на српском језику. Омогућава напредну детекцију говорних сегмената на мобилном телефону и другим уређајима који поседују одговарајуће техничке карактеристике, без обзира на релативно ниску процесорску моћ таквих уређаја. Техничко решење доприноси поузданом препознавању корисничких команди, захтева и упита (нпр. позивање, слање порука, промена поставки уређаја, измена информација о контактима и слично), уз високе перформансе по питању брзине и тачности, како у стандардном окружењу, тако и у условима повећаног нивоа позадинског шума (нпр. музика, разговор у кафићу, канцеларијска бука, саобраћајна бука и слично).

Докази (прилози):

- Писано мишљење два рецензента - експерта из области техничког решења.
- Документ којим се доказује да предузеће *АлфаНум* користи ово техничко решење у развоју система говорних технологија.
- Документ којим се доказује да предузеће *Ахон - интелигентни системи* користи ово техничко решење у развоју система говорних технологија.

Документација за техничко решење: Напредна детекција говорних сегмената на
мобилном телефону применом микрофонског низа и декодера

- Уверење о признавању техничког решења Наставно-научног већа Факултета техничких наука, на основу мишљења рецензената и приложених доказа, издато након одговарајуће процедуре на Департману за енергетику, електронику и телекомуникације.

Нови Сад, октобар 2016. године.

Подносилац пријаве

Проф. др Владо Делић
Руководилац пројекта ТР32035

POTVRDA

Ovim potvrđujemo da je u preduzeću AlfaNum DOO Novi Sad počev od 01.07.2016. godine počelo da se koristi tehničko rešenje *Napredna detekcija govornih segmenata na mobilnom telefonu primenom mikrofonskog niza i dekodera*, koje je zajednički razvijeno od strane Fakulteta tehničkih nauka u Novom Sadu i preduzeća AlfaNum, a koristi se za potrebe preduzeća AlfaNum i AXON DOO.

U Novom Sadu,

01.10.2016.

ALFANUM

Darko Pekar

Darko Pekar



POTVRDA

Ovim potvrđujemo da je u preduzeću AXON DOO počev od 01.07.2016. godine počelo da se koristi tehničko rešenje *Napredna detekcija govornih segmenata na mobilnom telefonu primenom mikrofonskog niza i dekodera*, koje je razvijeno od strane AlfaNum DOO Novi Sad, Bulevar Evrope 38a Novi Sad, u okviru projekta tehnološkog razvoja „Razvoj dijaloških sistema za srpski i druge jezike“.

U Novom Sadu,

01.10.2016.



Direktor,

Vesna Rehner



Račun br. 095a/16

Kupac

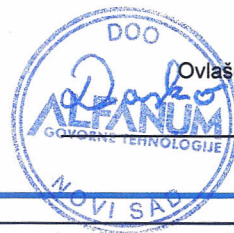
Naziv: AXON DOO
Adresa: Polgar Andraša 38a
Mesto: Novi Sad
P.I.B: 107424643

Datum: 30.jun.16
Mesto: Novi Sad
Datum PDO: 30.jun.16

Br.art.	Opis	Količina	Jed. mere	Cena bez PDV (din.)	Rabat	Por. Osnova (din.)	Stopa PDV	PDV (din.)	Ukupno (din.)
1	Modul za detekciju govornih segmenata na mobilnom telefonu primenom mikrofonskog niza i dekodera"					50.000,00	20%	10.000,00	60.000,00

Ukupno bez PDV-a	50.000,00
Poreska osnovica	50.000,00
PDV iznos	10.000,00
Za uplatu	60.000,00

Slovima: šezdeset hiljada dinara 00/100



Ovlašćeno lice



УНИВЕРЗИТЕТ
У НОВОМ САДУ



ФАКУЛТЕТ
ТЕХНИЧКИХ НАУКА

Трг Доситеја Обрадовића 6, 21000 Нови Сад, Република Србија
Деканат: 021 6350-413; 021 450-810; Централа: 021 485 2000
Рачуноводство: 021 458-220; Студентска служба: 021 6350-763
Телефакс: 021 458-133; e-mail: ftndean@uns.ac.rs

ИНТЕГРИСАНИ
СИСТЕМ
МЕНАџМЕНТА
СЕРТИФИКОВАН ОД:



Наш број: 01-сл

Ваш број:

Датум: 2016-11-03

ИЗВОД ИЗ ЗАПИСНИКА

Наставно-научно веће Факултета техничких наука у Новом Саду, на 25. редовној седници одржаној дана 26.10.2016. године, донело је следећу одлуку:

-непотребно изостављено-

ТАЧКА 10. Питања научноистраживачког рада и међународне сарадње

Тачка 10.2.1: У циљу верификације новог техничког решења усвајају се рецензенти:

- Доц. др Бранислав Геразов, Факултет за електротехнику и информационе технологије, Скопље, Македонија
- Др Јернеја Жганец, развојни саветник, Алпинеон д.о.о. Љубљана, Словенија

Назив техничког решења:

**"НАПРЕДНА ДЕТАКЦИЈА ГОВОРНИХ СЕГМЕНАТА НА МОБИЛНОМ ТЕЛЕФОНУ
ПРИМЕНОМ МИКРОФОНСКОГ НИЗА И ДЕКОДЕРА"**

Аутори техничког решења: Бранислав Поповић, Едвин Пакоци, Дарко Пекар, Владо Делић.

-непотребно изостављено-

Записник водила:

Јасмина Димић, дипл. правник

Тачност података оверава:
Секретар

Иван Нешковић, дипл. правник

Декан

Проф. др Раде Дорословачки



РЕЦЕНЗИЈА ТЕХНИЧКОГ РЕШЕЊА

Подаци о техничком решењу:

Назив техничког решења:	Напредна детекција говорних сегмената на мобилном телефону применом микрофонског низа и декодера
Аутори техничког решења:	Бранислав Поповић, Едвин Пакоци, Дарко Пекар, Владо Делић
Реализатори:	ФТН и АлфаНум д.о.о. из Новог Сада
Пројекти на којим је развијено:	“Развој дијалошких система за српски и друге јужно-словенске језике”, ТР32035 код МПНТР (2011-2016)
Област на коју се односи:	Електроника, телекомуникације и информационе технологије
Корисници:	Предузећа АлфаНум д.о.о. и Ахон - интелигентни системи из Новог Сада, од 2016. године
Категорија техничког решења:	Ново техничко решење (метода) примењено на националном нивоу (М82)

Подаци о рецензенту:

Име, презиме и звање:	Доц. др Бранислав Геразов
Ужа научна област за коју је изабран у звање, датум избора у звање и назив факултета:	Изабран у звање <i>доцент</i> 16.06.2016. на <i>Факултету за електротехнику и информационе технологије</i> , Скопље, Македонија, за у.н.о. <i>електроника</i>
Установа где је запослен:	Факултет за електротехнику и информационе технологије, Скопље, Македонија

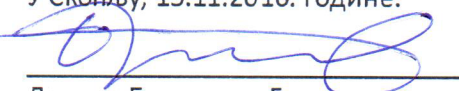
Стручно мишљење рецензента:

Резултат научно-истраживачког рада представља **ново техничко решење (методу) примењено на националном нивоу (М82)** у смислу Правилника о поступку, начину вредновања и квантитавном исказивању научноистраживачких резултата истраживача, донетог од стране министра просвете, науке и технолошког развоја на основу члана 70. ст. 7. и 8. и члана 86. став 2. Закона о научноистраживачкој делатности („Службени гласник РС”, бр. 110/05, 50/06-исправка, 18/10 и 112/15).

Образложење за техничко решење (ТР):

- ТР представља нову методу детекције говорне активности приликом употребе мобилног телефона. Користи се у оквиру апликације за управљање телефоном на српском језику.
- ТР је веома значајно у системима који користе говорне технологије, посебно у условима повишене позадинске буке. Комбинацијом неколико метода за детекцију говора у стерео аудио сигналу добијен је изузетно робустан систем високих перформанси.
- Развој ТР пратио је већи број научних и стручних радова који су пријављивани као резултати на пројектима МПНТР.
- ТР се успешно примењује у предузећима АлфаНум и Ахон – интелигентни системи у Новом Саду. Има могућност примене у многим системима који укључују препознавање говора.

У Скопљу, 15.11.2016. године.


Доц. др Бранислав Геразов

RECENZIJIA TEHNIČKOG REŠENJA

Podaci o tehničkom rešenju:

Naziv tehničkog rešenja:	Napredna detekcija govornih segmenata na mobilnom telefonu primenom mikrofonskog niza i dekodera
Autori tehničkog rešenja:	Branislav Popović, Edvin Pakoci, Darko Pekar, Vlado Delić
Realizatori:	FTN i AlfaNum d.o.o. iz Novog Sada
Projekti na kojim je razvijeno:	“Razvoj dijaloških sistema za srpski i druge južno-slovenske jezike”, TR32035 kod MPNTR (2011-2016)
Oblast na koju se odnosi:	Elektronika, telekomunikacije i informacione tehnologije
Korisnici:	Preduzeća AlfaNum d.o.o. i Axon – inteligentni sistemi d.o.o. iz Novog Sada, od 2016. godine
Kategorija tehničkog rešenja:	Novo tehničko rešenje (metoda) primenjeno na nacionalnom nivou (M82)

Podaci o recenzentu:

Ime, prezime i zvanje:	dr Jerneja Žganec Gros, <i>razvojni savetnik</i>
Uža naučna oblast za koju je izabran u zvanje, datum izbora u zvanje i naziv fakulteta:	Izabrana u zvanje <i>razvojni savetnik za IKT</i> 26.05.2003. godine od strane <i>Komisije za razvrščanje v raziskovalne nazive za naučnu oblast elektrotehnika</i> .
Ustanova gde je zaposlen:	Alpineon razvoj in raziskave, d.o.o., Ljubljana

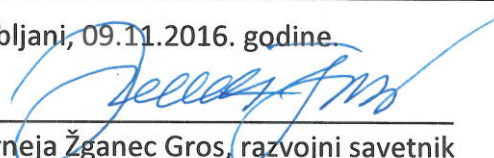
Stručno mišljenje recenzenta:

Rezultat naučno-istraživačkog rada “Napredna detekcija govornih segmenata na mobilnom telefonu primenom mikrofonskog niza i dekodera” predstavlja **ново tehničko rešenje (metodu) primenjeno na nacionalnom nivou (M82)** u smislu Pravilnika o postupku, načinu vrednovanja i kvantitativnom iskazivanju naučnoistraživačkih rezultata istraživača, donetog od strane ministra prosvete, nauke i tehnološkog razvoja na osnovu člana 70. st. 7. i 8. i člana 86. stav 2. Zakona o naučnoistraživačkoj delatnosti („Službeni glasnik RS”, br. 110/05, 50/06-ispravka, 18/10 i 112/15).

Obrazloženje za tehničko rešenje (TR):

- TR je nova metoda detekcije govornog segmenta u aplikacijama za zadavanje govornih komandi mobilnim telefonima koji koriste Android operativni sistem, na srpskom.
- TR omogućuje visoke performanse sistema, uključujući efikasan rad, pouzdanu detekciju govora tamo gde ga ima, kao i prepoznavanje tišine, odnosno šuma tamo gde govora nema. Dobro radi čak i slučaju prilično visokog nivoa pozadinske buke bilo koje vrste.
- TR je zasnovano na primeni softverskih rešenja i naprednih govornih tehnologija za srpski jezik, razvijenim korišćenjem naučnih metoda na projektima MPNTR.
- TR se primenjuje u preduzećima AlfaNum i Axon – inteligentni sistemi u Novom Sadu, a može se ugraditi i u druge sisteme koji uključuju ulazni govorni signal (komande).

U Ljubljani, 09.11.2016. godine.


dr Jerneja Žganec Gros, *razvojni savetnik*



УНИВЕРЗИТЕТ
У НОВОМ САДУ



ФАКУЛТЕТ
ТЕХНИЧКИХ НАУКА

Трг Доситеја Обрадовића 6, 21000 Нови Сад, Република Србија
Деканат: 021 6350-413; 021 450-810; Централа: 021 485 2000
Рачуноводство: 021 458-220; Студентска служба: 021 6350-763
Телефакс: 021 458-133; e-mail: ftndean@uns.ac.rs

ИНТЕГРИСАНИ
СИСТЕМ
МЕНАџМЕНТА
СЕРТИФИКОВАН ОД:



Наш број: 01.сл

Ваш број:

Датум: 2016-12-07

ИЗВОД ИЗ ЗАПИСНИКА

Наставно-научно веће Факултета техничких наука у Новом Саду, на 26. редовној седници одржаној дана 30.11.2016. године, донело је следећу одлуку:

-непотребно изостављено-

ТАЧКА 11. Питања научноистраживачког рада и међународне сарадње

Тачка 11.1.: На основу позитивног извештаја рецензената верификује се техничко решење (М82) под називом:

"НАПРЕДНА ДЕТАКЦИЈА ГОВОРНИХ СЕГМЕНАТА НА МОБИЛНОМ ТЕЛЕФОНУ ПРИМЕНОМ МИКРОФОНСКОГ НИЗА И ДЕКОДЕРА"

Аутори техничког решења: Бранислав Поповић, Едвин Пакоци, Дарко Пекар, Владо Делић.

-непотребно изостављено-

Записник водила:

Јасмина Димић, дипл. правник

Тачност података оверава:
Секретар

Иван Нешковић, дипл. правник

Декан

Проф. др Раде Дорословачки

