Образац за пријаву техничког решења[1]

Назив техничког решења	Побољшање перформанси система за гласовно позивање (VoiceDial) применом Chain декодера
Аутори техничког решења	Бранислав Поповић, Едвин Пакоци, Дарко Пекар, Владо Делић
Категорија техничког решења	Битно побољшано техничко решење на националном нивоу (М84)

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

Техничко решење је реализовано у оквиру пројекта технолошког развоја "Развој дијалошких система за српски и друге јужнословенске језике" (ТР32035, 2011-2017) на Факултету техничких наука и у предузећу АлфаНум у Новом Саду.

Ко користи техничко решење:

Предузеће АлфаНум у Новом Саду, као партиципант на пројекту ТР32035.

Година када је техничко решење урађено:

Развој система за гласовно позивање (VoiceDial) отпочео је 2010. године, док је развој говорних технологија примењених у склопу техничког решења, уз прикупљање говорних база за обуку и тестирање, текао у континуитету, у периоду од 2005. до 2017. године, у оквиру технолошких пројеката МПНТР. Развој битно побољшане верзије техничког решења, засноване на примени Chain декодера, почео је крајем 2016. године и завршен је 2017. године.

Ко је прихватио-примењује техничко решење:

Решење је иницијално реализовано за потребе развоја говорних технологија у предузећу АлфаНум, а користи се и у развојном тиму на Факултету техничких наука у Новом Саду. Унапређена верзија техничког решења примењује се од стране предузећа Axon - интелигентни системи у Новом Саду и Сага д.о.о. у Београду од 2017. године.

[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. 24/2016 и 21/2017).

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, које га је увело у свој производни програм. Побољшање је прихваћено од стране предузећа Axon - интелигентни системи у Новом Саду и Сага д.о.о. у Београду од 2017. године, док је основна верзија у примени од 2011. године и представља први такав производ примењен на јужнословенском говорном подручју.

2) Техничко решење представља битно побољшање постојећег техничког решења на националном нивоу (М84), реализовано у оквиру пројекта ТР32035:

   - *Дарко Пекар, Драган Кнежевић, Милан Сечујски, Никша Јаковљевић, Наташа Вујновић Седлар, Владо Делић, „Систем за гласовно позивање (VoiceDial)“, Факултет техничких наука и Алфанум, 2012. [М82]*

   Засновано је на већем броју публикација, верификованих техничких и развојних решења која су пријављивана као резултати на технолошким и иновационим пројектима МПНТР у периоду од 2005. до 2017. године. Поједини елементи техничког решења презентовани су на више домаћих и међународних научних скупова, на пример:

   - *Едвин Пакоци, Бранислав Поповић, Дарко Пекар, „Fast Sequence-Trained Deep Neural Network Models for Serbian Speech Recognition“, Дигитална обрада говора и слике, ДОГС 2017, Нови Сад, Србија, 22-25. новембар 2017, pp. 25-28, ISBN: 978-86-7892-993-9.*

   - *Бранислав Поповић, Едвин Пакоци, Дарко Пекар, „End-to-End Large Vocabulary Speech Recognition for the Serbian Language“, Proc. of the 19th SPECOM, Speech and Computer, Хатфилд, Хартфордшир, УК, 12-16. септембар 2017, Alexey Karpov, Iosif Mporas, Rodmonga Potapova (Eds.), Lecture Notes in Artificial Intelligence (LNAI), Springer, vol. 10458, pp. 343-352, DOI: 10.1007/978-3-319-66429-3_33, Series ISSN: 0302-9743, Electronic ISSN: 1611-3349, Print ISBN: 978-3-319-66428-6, Online ISBN: 978-3-319-66429-3.*

   - *Едвин Пакоци, Бранислав Поповић, Дарко Пекар, „Language Model Optimization for a Deep Neural Network Based Speech Recognition System for Serbian“, Proc. of the 19th SPECOM, Speech and Computer, Хатфилд, Хартфордшир, УК, 12-16. септембар 2017, Alexey Karpov, Iosif Mporas, Rodmonga Potapova (Eds.), Lecture Notes in Artificial Intelligencee (LNAI), Springer, vol. 10458, pp. 483-492, DOI: 10.1007/978-3-319-66429-3_48, Series ISSN: 0302-9743, Electronic ISSN: 1611-3349, Print ISBN: 978-3-319-66428-6, Online ISBN: 978-3-319-66429-3.*

3) Дато је писано мишљење два рецензента - експерта из области техничког решења:

   - Др Мишко Суботић, Центар за унапређење животних активности (Београд, Србија)

   - Проф. др Јернеја Жганец Грос, *Alpineon razvoj in raziskave, d.o.o.* (Љубљана, Словенија)

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 битно побољшано техничко решење на националном нивоу (М84), у складу са Правилником Министарства.

На који начин се користи (кратак опис):

Техничко решење представља битно побољшање система за гласовно позивање (VoiceDial), остварено применом Chain декодера. За разлику од основне верзије решења, развијене за потребе великих и средњих предузећа (од неколико десетина до неколико хиљада корисника), побољшана верзија подржава знатно веће речнике (нова верзија декодера тестирана је на великим речницима, са преко 120000 речи, док је основна верзија омогућавала задовољавајући ниво тачности на речницима не већим од неколико хиљада речи). Нова верзија модула за аутоматско препознавање говора (*Automatic Speech Recognition* - ASR) пружа знатно побољшање тачности препознавања у односу на првобитну верзију, чак и у случајевима високо зашумљених сигнала, и осигурава стабилан резултат, уз релевантне мере поузданости препознавања по изговореној речи (име и презиме контакта). У поређењу са другим декодерима заснованим на примени неуралних мрежа, применом Chain декодера остварена је слична или повећана тачност препознавања, уз знатно снижено време декодовања (повећану брзину одзива). Примена је непромењена у односу на претходну верзију - у комуникацији са постојећом телефонском централом врши се трансфер позива на одговарајући локал унутар или изван организације. Избор телефонског броја (контакта) врши се путем гласа. Позив се прослеђује говорном аутомату и успоставља аутоматски, по успешном препознавању. Контакти се уносе у систем путем специјално дизајнираног веб интерфејса, чиме је елиминисана потреба за инсталацијом додатног софтвера и ангажовањем људског оператера. Обезбеђена је заштита система од злоупотреба, применом корисничког имена и лозинке и различитих нивоа приступа (корисник, супервизор и администратор).

Опис техничког решења: *Побољшање перформанси система за гласовно позивање (VoiceDial) применом Chain декодера*

Битно побољшање техничког решења на националном нивоу остварено је применом говорних технологија развијеним на пројектима Министарства, коришћењем научних метода, и пропраћено већим бројем научних и стручних радова објављеним у протеклих годину дана (резултати пројекта ТР32035). Основна верзија система за гласовно позивање представља прво такво техничко решење на подручју јужнословенских језика, прихваћено од стране неколико највећих систем интегратора

из Београда, док је побољшана верзија система већ прихваћена од стране компанија Axon - интелигентни системи у Новом Саду и Сага д.о.о. у Београду, непосредно након имплементације, средином 2017. године.

Област на коју се техничко решење односи:

Техничко решење припада области информационо-комуникационих технологија (ICT), и представља спој већег броја научних и техничких дисциплина - вештачке интелигенције, машинског учења, обраде говорних сигнала, дигиталне обрада звука, препознавања и синтезе говора, дијалошких система и софтверског инжењеринга.

Проблем који се техничким решењем решава:

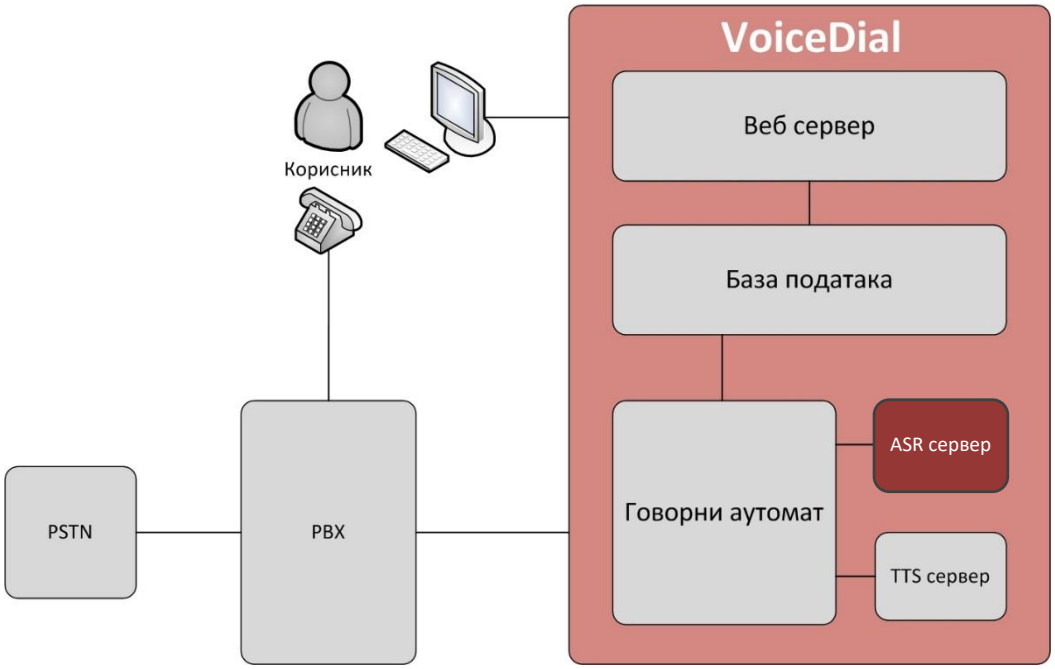
Интеграцијом усавршеног модула за препознавање говора остварена је аутоматизација процеса позивања у оквиру једног предузећа, уз битно побољшање тачности препознавања и прихватљиво време одзива, чак и у условима повећаног нивоа шума (канцеларијска, саобраћајна и друге врсте буке). Контакти се позивају по називу одељења, имену и презимену, изговарањем назива, само имена, само презимена, или пуног имена и презимена, како за интерне позиве (унутар фирме), тако и у случају спољног позиваоца. У сваком тренутку, применом поузданих мера тачности препознавања по речи, систем ће извршити селекцију тражених контаката и бројева телефона. По извршеној селекцији, кориснику ће применом система за претварање текста у говор (*Text-to-Speech*, TTS) бити понуђене све расположиве алтернативе, при чему је сваком контакту могуће придружити већи број бројева телефона, док дуплирање уноса није подржано. Систем се може користити и за добијање броја или бројева телефона придружених одређеном контакту из интерне базе података. Контакти се у систем уносе ручно, или се могу учитати из екстерне датотеке. На тај начин аутоматозавана је и поједностављена процедура позивања, уз битно побољшану поузданост система и капацитет уноса у односу на претходну верзију.

Стање решености тог проблема у свету:

Систем за гласовно позивање (VoiceDial) убраја се у групу производа заснованих на примени говорних технологија, попут производа компаније Nuance Communications, *CMG Speech* предузећа Aastra Technologies Limited, *Avaya one-X Speech* предузећа Avaya, и њима сличних производа намењених Андроид и iOS оперативним системима. Имплементација ових производа условљена је степеном развоја технологија за препознавање и синтезу говора за одређени језик, на које се ови системи ослањају. Предложено техничко решење користи говорне технологије посебно развијене и прилагођене српском језику и другим јужнословенским језицима. Примењени су најсавременији модели за аутоматско препознавање говора, специјално обучени за сваки од понуђених језика и прилагођени препознавању како у нормалним, тако и у условима повишеног нивоа буке.

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

Архитектура система приказана је на Слици 1. Основну функционалну јединицу система чини говорни аутомат, у оквиру ког је имплементирана логика комуникације са корисником. Упити корисника прослеђују се ASR серверу на препознавање, док се одговори кориснику генеришу применом TTS модула, у комуникацији са интерном базом података. База садржи информације о корисницима система и њиховом праву приступа, као и комплетан именик генерисан од стране корисника. Приступ је омогућен путем посебног линка, док се пријава на систем обавља укуцавањем корисничког имена и лозинке.



Слика 1. Архитектура VoiceDial система са назначеном ASR компонентом чије је побољшање извршено у склопу техничког решења

Контакти се у систем могу додавати ручно или аутоматски - уносом контаката из екстерне CSV датотеке одговарајуће структуре, при чему се изговор имена контакта посебно записује у пољу за изговор, и може се разликовати од стварно наведеног имена и презима. Деле се на администраторске и корисничке. Администраторски контакти су стандардно видљиви свим запосленима у оквиру предузећа (опционо) и њихова измена од стране корисника није могућа. Кориснички контакти имају предност у односу на администраторске контакте и дефинишу се од стране корисника система.

Битно побољшање система за аутоматско позивање (VoiceDial) остварено је заменом ASR компоненте, односно применом Chain декодера и модула специјално обучених за одговарајући језик апликације и различите нивое шума. Поред подршке за знатно

већи број уноса (преко 10 пута) и значајног побољшања по питању тачности препознавања уз задовољавајућу брзину одзива, извршена је имплементација одговарајућих мера поузданости, чиме је обезбеђен стабилан и поуздан рад система у свим условима.

Chain модели спадају у групу модела заснованих на примени дубоких неуралних мрежа и скривених Марковљевих модела (DNN-HMM). Обука Chain модела врши се секвенцијално, применом логаритамске вероватноће исправне секвенце (за сваку секвенцу у бази за обуку) као функције циља. Заснована је на примени метода максималне узајамне информације (*Maximum Mutual Information*, MMI), при чему се на излазу неуралне мреже користи 3 пута умањен број фрејмова, што доприноси значајном смањењу рачунске сложености и олакшава декодовање у реалном времену (3 пута брже декодовање уз повећану тачност у односу на стандардне DNN моделе). Услед снижене брзине одабирања, користи се посебан поједностављен тип HMM модела који дозвољава пролазак кроз цео HMM у једном фрејму (постоји само једно емитујуће стање), уз фиксиране вероватноће останка у стању и прелаза у следеће стање (0,5). При томе, при првом фрејму кроз дати HMM емитује се симбол који означава расподелу вероватноћа везану за њега, док се за све остале фрејмове емитује исти специјални бланко симбол, што омогућава појаву секвенци попут *a*, *ab*, *abb*, *abbb* и слично, где је *b* ознака за бланко знак. Улазна поравнања секвенци за обуку на нивоу фрејма конвертована су у складу са редукованом брзином одабирања, и као таква се користе за изградњу новог контекстног стабла за кластеризацију HMM стања, док је процедура кластеровања иста као и код стандардних GMM модела.

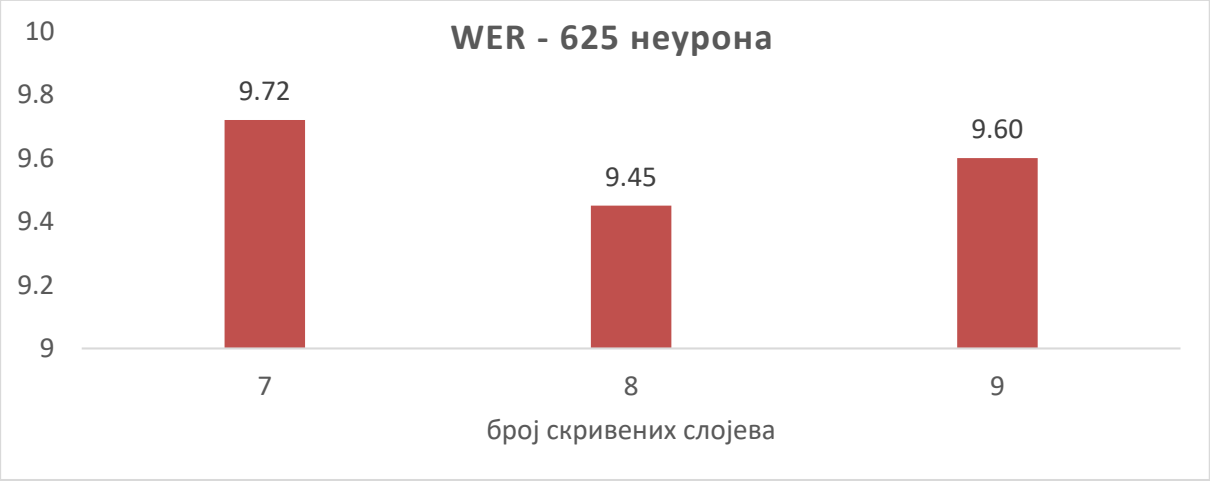
Функцију акустичког модела врше временски закашњене неуралне мреже (*Time-Delay Neural Networks* - TDNN), које се обучавају применом критеријума унакрсне ентропије (*cross-entropy*). Почетни слојеви мреже узимају у обзир по 3 узастопна фрејма (претходни - *t-1*, тренутни - *t* и наредни - *t+1*), док они најскривенији слојеви такође користе по 3 фрејма, али одвојена за по 3 фрејма (*t-3*, *t*, *t+3*), што омогућава њихову евалуацију тек при сваком трећем фрејму. Процедура обуке подразумева прерачунавање бројиоца (нумератора - који одговара вероватноћи дате исправне секвенце) и имениоца (деноминатора - који енкодује све дозвољене секвенце, односно одговара укупној вероватноћи свих дозвољених секвенци) функције циља и њиховог међусобног односа (који треба максимизовати), при чему не постоји потреба за нормализацијом DNN излаза. Сложенија израчунавања у именионцу функције циља се врше применом графичког процесора (GPU) и *forward-backward* алгоритма, чиме се искоришћава синхронизован приступ меморији између језгара. Једноставнија израчунавања у бројиоцу функције циља врше се на централној процесорској јединици (CPU). Брзина одабирања је постављена на 100 фрејмова по секунди.

Нумератор и деноминатор се чувају у формату аутомата са коначним бројем стања (*Finite State Acceptor* - FSA), са ознакама лабела (симбола) уместо стања на прелазима.

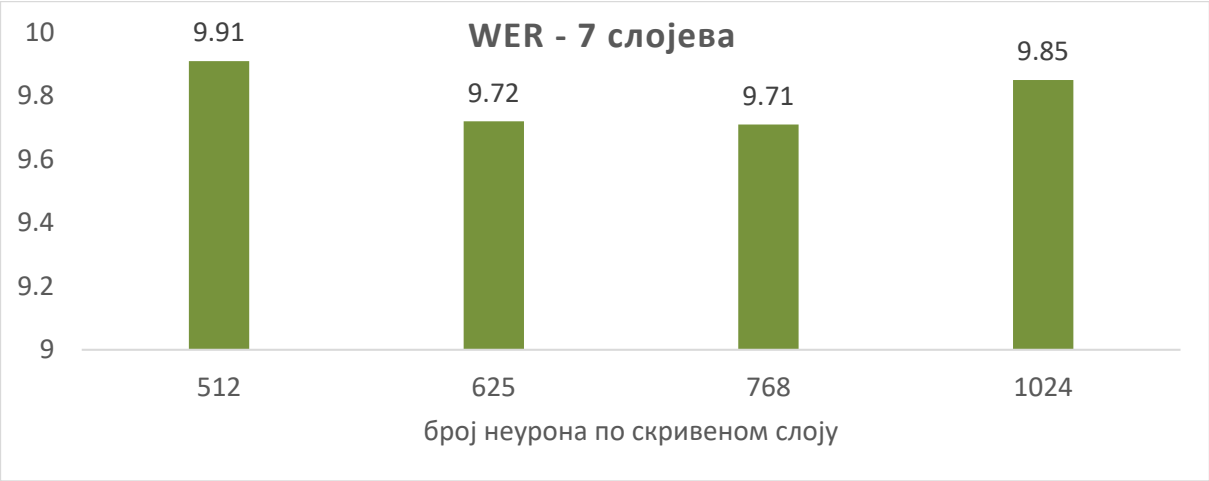
За обуку деноминаторског графа користи се специјални поједностављени језички модел на нивоу фонема (4-грам), при чему се ублажавање и одсецање врши само за и изнад 3-грам нивоа. У циљу даље редукције укупног броја стања и транзиција, примењује се посебна процедура детерминизације и минимизације. У случају нумераторског графа, улазна поравнања на нивоу фрејма претварају се у мреже које представљају све дозвољене алтернативне изговоре задате реченице, након чега се врши конверзија у граф на нивоу фонема, и коначно у реченични FSA. Свако FSA стање се може идентификовати индексом фрејма, што је корисно при паралелизацији израчунавања на нумератору. Коначни нумераторски граф увек садржи подскуп путања садржаних у деноминаторском графу.

У сврху тестирања, примењен је 3-грам језички модел трениран на скупу за обуку и додатним подацима из српског новинског текстуалног корпуса. Коначан модел језика садржи преко 120000 речи (приближно 1,3 милиона 2-грама и 358000 3-грама, уз вредност прага одсецања постављену на 10^{-7}). База садржи преко 200 сати аудио снимака за обуку за различите нивое шума, од чега је нешто мање од 20 сати коришћено за потребе тестирања. Извршена је оптимизација броја скривених слојева неуралне мреже (7 до 9), броја неурона по једном слоју (512 до 1024), броја итерација обуке (3 до 5), комплексности HMM структуре (3000 или 4000 стања), као и додатних параметара (*Пакоци, Поповић, Пекар, ДОГС 2017*).

На Слици 2 приказани су резултати тестирања у виду стопе грешке препознавања по речи (*Word Error Rate* - WER) при промени броја скривених слојева.

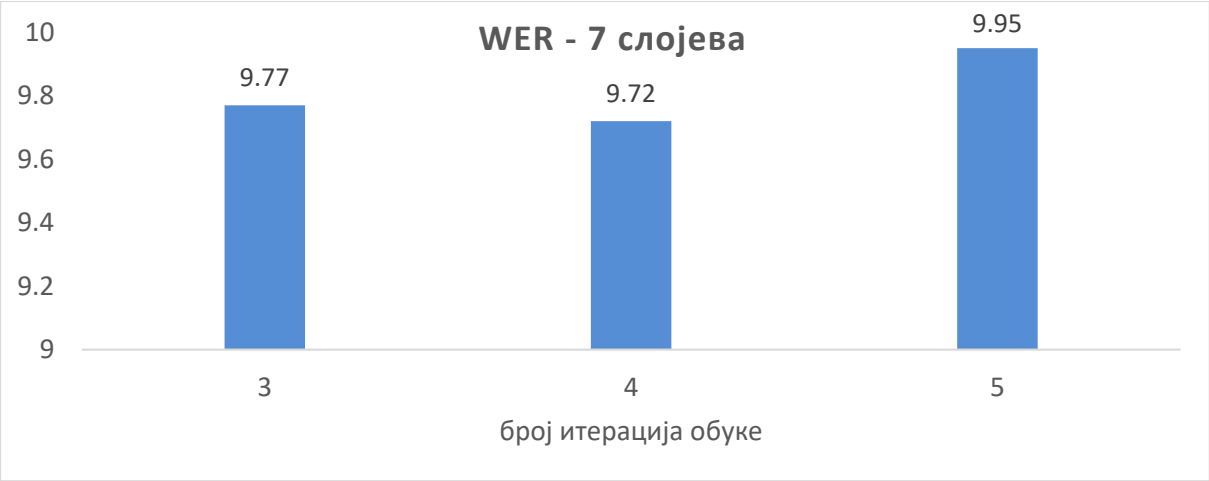


Слика 2. Грешка препознавања по речи при промени броја скривених слојева



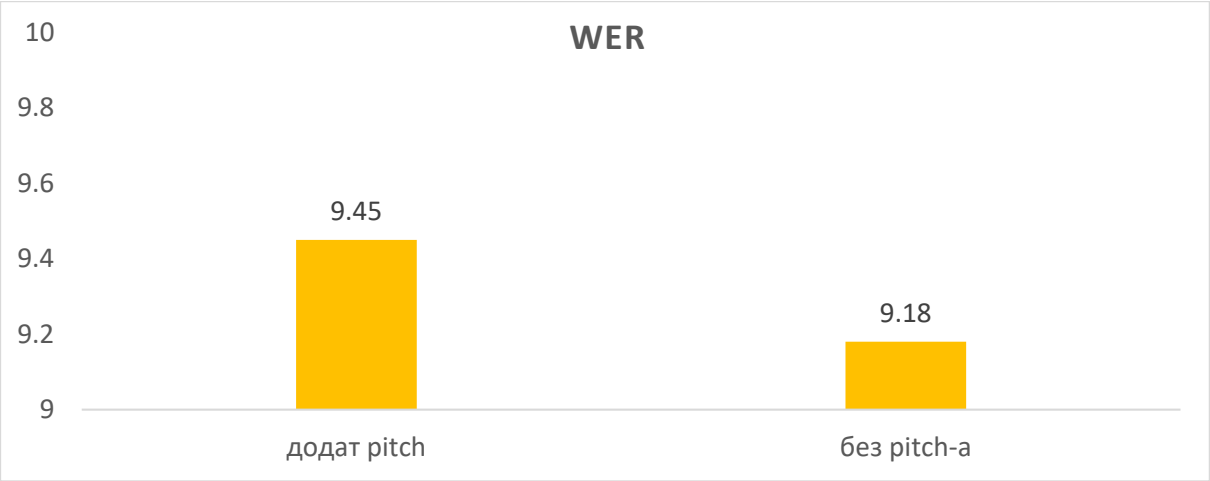
Слика 3. Грешка препознавања по речи при промени броја неурона по скривеном слоју

Слика 3 приказује резултате тестирања при промени броја неурона по једном скривеном слоју, док су на Слици 4 представљени резултати тестирања при промени броја итерација обуке.



Слика 4. Грешка препознавања по речи при промени броја итерација обуке

Коначно, на Слици 5 су представљени резултати са и без примене додатних *pitch* обележја (три обележја везана за основну учестаност говорног сигнала). Примери 2 до 5 су приказани за српски језик (резултати су слични и за преостале језике).



Слика 5. Грешка препознавања по речи при употреби додатног *pitch* обележја

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

Поред стандардних опција, попут интеграције са постојећом телефонском централом, уноса контаката из постојећег телефонског именика, персонализације именика и приступа систему путем Интернета, битно побољшана верзија система нуди следеће могућности:

- Стабилно и поуздано препознавање изговореног назива контакта, како у нормалним условима, тако и у условима повишеног нивоа шума.
- Ефикасно препознавање великог броја контаката у реалном времену и са задовољавајућом тачношћу (преко 10 пута више у односу на основну верзију) и систем позивања 24/7, без потребе за ангажовањем људских оператера.
- Разликовање назива контаката који слично звуче применом акустичке сличности и релевантних мера поузданости по свакој изговорених речи. Поменуте мере пружају основ за додатно усавршавање система, применом одговарајућих пенала и прагова.
- Аутоматско ажурирање именика и корисничких граматика при сваком уносу.
- Посебно обучени акустички модели за српски и друге јужнословенске језике.

Докази (прилози):

- Писано мишљење два рецензента - експерта из области техничког решења.
- Докази о примени техничког решења у предузећима АлфаНум и Axon - интелигентни системи у Новом Саду, и Сага д.о.о. у Београду.
- 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, децембар 2017. године.

Подносилац пријаве

Проф. др Владо Делић
Руководилац пројекта ТР32035

AlfaNum d.o.o., Novi Sad
Bulevar vojvode Stepa 40
021/4750204
PIB: 102948614
Mat.br. 08790949
Tekući račun: 340-11002697-66
Šifra delatnosti: 7219

Novi Sad
06.12.2017.

POTVRDA

Ovim potvrđujemo da je 01.07.2017. godine u preduzeću AlfaNum d.o.o. počelo da se koristi tehničko rešenje *Poboljšanje performansi sistema za glasovno pozivanje (VoiceDial) primenom Chain dekodera*, koje je zajednički razvijeno od strane Fakulteta tehničkih nauka Novi Sad i preduzeća Alfanum i koristi se za potrebe internog testiranja i daljeg razvoja rešenja. Rešenje je inicijalno instalirano za potrebe testiranja, ali ga preduzeće AlfaNum koristi i u promociji govornih tehnologija za srpski jezik.

U ime firme:

Direktor
Darko Pekar, dipl.ing.

Ovlašćeno lice

Darko Pekar

DOO ALFANUM GOVORNE TEHNOLOGIJE NOVI SAD

axon inteligentni sistemi

Bulevar vojvode Stepe br. 40, 21000 Novi Sad, Srbija
Web: www.axon-is.co.rs

POTVRDA

Ovim potvrđujemo da je u julu mesecu 2017. godine u preduzecu Axon - Inteligentni sistemi d.o.o. instalirano tehničko rešenje - Poboljšanje performansi sistema za glasovno pozivanje (VoiceDial) primenom Chain dekodera, kompanije AlfaNum i da je od tada u upotrebi.
Sistem nesmetano funkcionise na preko 10 lokala. VoiceDial sistem je inicijalno instaliran za potrebe testiranja i promocije, ali ga preduzece Axon koristi i za sopstvene potrebe.

U Novom Sadu 11.12.2017.

axon inteligentni sistemi doo novi sad

SAGA
new frontier group

Privatno društvo sa ograničenom odgovornošću za promet roba i usluga u zemlji i inostranstvu
»S A G A«
Broj 10387
13.12. 2017

POTVRDA

Ovim potvrđujemo da je jula 2017. godine u preduzecu Saga d.o.o. Beograd instalirano tehničko rešenje „Poboljšanje performansi sistema za glasovno pozivanje (VoiceDial) primenom Chain dekodera", kompanije AlfaNum i da je od tada u upotrebi.

Sistem nesmetano funkcionise na preko 100 lokala. VoiceDial sistem je inicijalno instaliran za potrebe testiranja i promocije, ali ga preduzece Saga koristi i za sopstvene potrebe.

Beograd, 12.12.2017.godine

Saga d.o.o. Beograd

PRIVATNO DRUŠTVO SA OGRANIČENOM ODGOVORNOŠĆU za promet roba i usluga u zemlji i inostranstvu SAGA Beograd XI

Saga d.o.o. Beograd
Bulevar Zorana Đinđića 64a, 11070 Beograd, Srbija
P +381 11 3108-500, F +381 11 3108-567, office@saga.rs
www.saga.rs
Privatno društvo sa ograničenom odgovornošću za promet robe i usluga u zemlji i inostranstvu Saga d.o.o. Beograd

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:
Ваш број:
Датум: 2017-10-26

ИЗВОД ИЗ ЗАПИСНИКА

Наставно-научно веће Факултета техничких наука у Новом Саду, на 44. редовној седници одржаној дана 25.10.2017. године, донело је следећу одлуку:

-непотребно изостављено-

Тачка 13.1. Верификација нових техничких решења и именовање рецензената

Тачка 13.1.3.: У циљу верификације новог техничког решења усвајају се рецензенти:

1. Др Мишко Суботић, научни сарадник, Центар за унапређење животних могућности
2. Проф. др Јернеја Жганец Гроц, развојни саветник, Alpineon razvoj in raziskave, Љубљана, Словенија

Назив техничког решења:

"ПОБОЉШАЊЕ ПЕРФОРМАНСИ СИСТЕМА ЗА ГЛАСОВНО ПОЗИВАЊЕ (VoiceDial) ПРИМЕНОМ Chain 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:	**Побољшање перформанси система за гласовно позивање (VoiceDial) применом Chain 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:	ТР32035 код МПНТР (2011-2017)
Област на коју се односи:	Електроника, телекомуникације и информационе технологије
Корисници:	Предузећа АлфаНум д.о.о. и Axon – интелигентни системи из Новог Сада, и Сага д.о.о. из Београда, од 2017. године
Категорија техничког решења:	Битно побољшано техничко решење на националном нивоу (М84)

Подаци о рецензенту:

Име, презиме и звање:	др Мишко Суботић, научни сарадник
Ужа научна област за коју је изабран у звање, датум избора у звање и назив факултета:	Изабран 25.03.2015. у звање научни сарадник на ЕТФ у Београду, у области техничко-технолошких наука - електроника, телеком. и информационе технологије
Установа где је запослен:	Центар за унапређење животних активности, Београд

Стручно мишљење рецензента:

Резултат научно-истраживачког рада "Побољшање перформанси система за гласовно позивање (VoiceDial) применом Chain декодера" представља **битно побољшано техничко решење на националном нивоу (М84)** 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. 24/2016 и 21/2017).

Образложење за техничко решење (ТР):

- Основна верзија ТР је први систем те врсте примењен за српски и друге јужнословенске језике, односно на одговарајућем говорном подручју.
- Побољшање ТР у виду коришћења модерног Chain декодера спада у групу најуспешнијих метода обуке и декодовања у савременом аутоматском препознавању говора.
- ТР је тестирано на речницима преко 120000 речи са запаженим резултатима.
- ТР омогућава препознавање у реалном времену без обзира на сложеност граматике и броја различитих речи за препознавање.
- ТР додатно омогућава и аутоматско ажурирање граматика (именика) за препознавање. Већ се успешно примењује у неколико фирми у Новом Саду и Београду.

У Београду, 27.12.2017. године.

Суботић Мишко

др Мишко Суботић

РЕЦЕНЗИЈА ТЕХНИЧКОГ РЕШЕЊА

Подаци о техничком решењу:

Назив техничког решења:	**Побољшање перформанси система за гласовно позивање (VoiceDial) применом Chain 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:	ТР32035 код МПНТР (2011-2017)
Област на коју се односи:	Електроника, телекомуникације и информационе технологије
Корисници:	Предузећа АлфаНум д.о.о. и Axon – интелигентни системи из Новог Сада, и Сага д.о.о. из Београда, од 2017. године
Категорија техничког решења:	Битно побољшано техничко решење на националном нивоу (М84)

Подаци о рецензенту:

Име, презиме и звање:	Проф. др Јернеја Жганец Грос, *razvojni savetnik*
Ужа научна област за коју је изабран у звање, датум избора у звање и назив факултета:	Изабрана у звање *profesor* 16.01.2017. године од стране *Fakulteta za industrijski inženiring Novo Mesto* за н.о. информациони системи. Изабрана у звање *razvojni savetnik za IKT* 26.05.2003. године од стране *Komisije za razvrščanje v raziskovalne nazive* за н.о. електротехника.
Установа где је запослен:	*Alpineon razvoj in raziskave, d.o.o.*, Љубљана, Словенија

Стручно мишљење рецензента:

Резултат научно-истраживачког рада "Побољшање перформанси система за гласовно позивање (VoiceDial) применом Chain декодера" представља **битно побољшано техничко решење на националном нивоу (М84)** 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. 24/2016 и 21/2017).

Образложење за техничко решење (ТР):

- ТР је унапређен систем за позивање путем гласа у фирмама, како интерно тако и спољно, поуздан и потпуно аутоматизован, за српски и друге јужнословенске језике.
- У односу на претходну верзију система, ТР доноси значајно повећан капацитет контаката, већу поузданост и поједностављену процедуру позивања.
- ТР добро ради чак и случају прилично високог нивоа позадинске буке било које врсте.
- ТР је засновано на примени софтверских решења и напредних говорних технологија за српски језик, развијеним коришћењем научних метода на пројектима МПНТР.
- ТР се већ успешно примењује у предузећима АлфаНум и Axon – интелигентни системи у Новом Саду, као и Сага у Београду.

У Љубљани, 27.12.2017. године.

Проф. др Јернеја Жганец Грос

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: 2018-01-03

ИЗВОД ИЗ ЗАПИСНИКА

Наставно-научно веће Факултета техничких наука у Новом Саду, на 49. редовној седници одржаној дана 27.12.2017. године, донело је следећу одлуку:

-непотребно изостављено-

ТАЧКА 13.1. Верификација нових техничких решења и именовање рецензената

Тачка 13.1.2.: На основу позитивног извештаја рецензената верификује се техничко решење под називом:

Назив техничког решења:

"ПОБОЉШАЊЕ ПЕРФОРМАНСИ СИСТЕМА ЗА ГЛАСОВНО ПОЗИВАЊЕ (VoiceDial) ПРИМЕНОМ Chain 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