Образац за пријаву техничког решења[1]

Назив техничког решења	**Декодер за препознавање континуалног говора на великим речницима**
Аутори техничког решења	**Никша Јаковљевић, Драгиша Мишковић, Едвин Пакоци, Роберт Мак, Бранислав Поповић**
Категорија техничког решења	**Ново техничко решење – прототип (М85)**

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

Техничко решење је реализовано у оквиру технолошког пројекта МПНТР "Развој дијалошких система за српски и друге јужнословенске језике" (ТР32035, 2011-14) 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 техничког решења започет је средином 2010. године, а довршен 2012. године. Техничко решење се од 2012. године примењује у оквиру предузећа АлфаНум.

Ко је прихватио-примењује техничко решење:

Техничко решење се примењује у оквиру предузећа АлфаНум и на Факултету техничких наука на Универзитету у Новом Саду, као и у фирми Аксон – интелигентни системи из Новог Сада.

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, а имплементирано и испитано на развојним системима у предузећима АлфаНум и Аксон, где се и даље користи.

[1] У складу са одредбама *Правилника о поступку и начину вредновања, и квантитавном исказивању научноистраживачких резултата истраживача*, који је 21.03.2008. године донео Национални савет за научни и технолошки развој Републике Србије («Службени гласник РС», бр. 38/2008).

2) Техничко решење је нови програмски систем базиран на више техничких и развојних решења која су пријављивана као резултати на технолошким и иновационим пројектима МПНТР у периоду од 2005. до 2012. године. Реализовано је у оквиру пројекта ТР32035 („Развој дијалошких система за српски и друге јужнословенске језике“). Користи се у развојним системима у предузећима АлфаНум и Аксон у Новом Саду.
3) Приложено је писано мишљење два рецензента, експерта из области.
    a. Проф. др Снежана Гудурић, Филозофски факултет у Новом Саду,
    b. Проф. др Слободан Јовичић, Електротехнички факултет у Београду.
4) Наставно-научно веће Департмана за енергетику, електронику и телекомуникације и Факултета техничких наука, на основу мишљења рецензената и приложених доказа, издали су Уверење о признавању техничког решења које потврђује да оно испуњава све услове да буде признато као техничко решење категорије М85 (ново техничко решење – прототип), у складу са Правилником Министарства.
5) Поједини елементи техничког решења презентовани су на неколико међународних научних скупова, на пример:
    - „A Decoder for Large Vocabulary Speech Recognition“, Никша Јаковљевић, Драгиша Мишковић, Марко Јанев, Дарко Пекар, *Proceedings of the 18th Int. Conference on Systems, Signals and Image Processing IWSSIP*, Сарајево, 2011.
    - „Једно решење декодера за аутоматско препознавање говора на великим речницима“, Никша Јаковљевић, Драгиша Мишковић, Марко Јанев, Дарко Пекар, *Зборник радова 18. Телеком. форума ТЕЛФОР*, Београд, 2010.

На који начин се користи (кратак опис):

Овај декодер за препознавање континуалног говора на великим речницима који броје неколико десетина хиљада речи користи се преко командне линије Windows оперативног система, позивом апликације декодера и прослеђивањем одговарајућих улазних параметара. Као улазне параметре очекује се да се наведу:

- Спецификације екстрактора обележја из аудио датотеке, у одговарајућем формату (SLib);
- Путања до датотеке која садржи скуп акустичких модела којима су моделовани говорни сегменти (монофони, бифони, трифони) – подржани су АлфаНум-ов формат, као и HTK формат записа акустичких модела;
- Путања до датотеке која садржи речник изговора речи (за све речи које могу да се препознају);
- Путања до датотеке која садржи модел језика на бази *N*-грама – датотека треба да буде у текстуалном ARPA формату;
- Путања до аудио датотека за које је потребно формирати транскрипције, тј. из којих треба добити излазну текстуалну секвенцу речи.

Опис техничког решења:
Декодер за препознавање континуалног говора на великим речницима

Овај декодер за препознавање континуалног говора на великим речницима је заснован на сопственим говорним технологијама и на пројектима Министарства науке и технолошког развоја, коришћењем научних метода, а без коришћења других патентних или лиценцних документација. Његов развој пратило је неколико научних и стручних радова, као и развојних и техничких решења, који су пријављивани као резултати на пројектима МПНТР у периоду 2005-2012. године. По својој архитектури и функционалностима које пружа овај декодер није потпуно јединствен, али је за разлику од већине реализован у објектно-орјентисаном програмском језику, а истовремено и солидно документован. Из те перспективе, као и узимајући у обзир да подржава и стандардне улазне формате (како текстуалне тако и бинарне), на њега се може гледати као систем са скупом карактеристика који ниједан други алат нема, поготово у окружењу.

Област на коју се техничко решење односи:

Техничко решење припада области Електроника, телекомуникације и информационе технологије, и као такво налази примену у неколико различитих области, попут телекомуникација и обраде сигнала, дигиталне обрада звука, аутоматског препознавања говора, дијалошких система.

Проблем који се техничким решењем решава:

Декодер за препознавање континуалног говора из великих речника представља основу за системе за препознавање континуалног говора који имају многоструку примену - у системима за диктирање и транскрипцију аудио снимака говора, као и у позивним центрима и дијалошким системима.

Овај декодер реализује основне технике које се користе у процесу препознавања континуалног говора на великим речницима. Иако се користи фонетски приступ моделовању говора, излаз декодера јесте највероватнија секвенца речи са одговарајућим временским границама између њих, са којима је много лакше радити (него са низом фонема). Речник се може прилагодити у зависности од конкретне примене – за задати домен може постићи одличне резултате. За разлику од препознавања на малим речницима, кључно је постојање ефикасних структура које убрзавају претраживање при декодовању и тиме омогућују препознавање у реалном времену. Декодер је намењен статистички моделованом говору, као и већина савремених система за препознавање говора. Због своје реализације у објектно-орјентисаном програмском језику C++ и тиме модуларне структуре, модификовање и

проширивање постојећег програмског кода је релативно једноставно. Целокупан код је сасвим солидно документован. Подржани су стандардни формати који се користе у препознавању говора (на пример HTK), тако да је могуће и коришћење овог декодера у комбинацији са другим доступним алатима.

Стање решености тог проблема у свету:

У последњих двадесетак година, за потребе аутоматског препознавања говора, у свету је развијено наколико јавно доступних алата, као што су на пример HTK (Hidden Markov Model Toolkit), Sphinx-4, Julius, Kaldi и RWTH ASR. Од њих, најпотпунију документацију има HTK, што у великој мери олакшава његово коришћење. Међутим, он је писан у програмском језику C, што за последицу има доста сложену структуру кода, чиме је значајно отежано његово проширивање и модификовање. Са друге стране, Sphinx-4, Kaldi и RWTH ASR су реализовани у објектно-орјентисаним програмским језицима (JAVA, C++ и опет C++, респективно) што омогућује једноставније мењање и проширивање постојећег кода, али је пратећа документација релативно скромна. За разлику од претходно наведених алата, који се састоје како од алата за обуку модела сегмената говора, тако и од алата за декодовање, Julius је алат који омогућава искључиво декодовање, а пошто подржава стандардне формате, може се користити у комбинацији са другим алатима, као што су HTK, CMU-Cam SLM (CMU-Cambridge Statistical Language Modeling Toolkit) и други.

Због ограничења претходно споменутих алата, као и немогућности њиховог коришћења у комерцијалне сврхе (лиценци), развијен је овај декодер на Факултету техничких наука у кооперацији са предузећем АлфаНум. Он је такође намењен модерном, статистичком начину моделовања говора, који омогућује рашчлањивање сложеног модела на простије подмоделе, од којих су најважнији акустички модел, модел језика и модел начина изговора. Декодер је имплементиран у програмском језику C++ и добро документован. Подржава стандардне формате и могуће га је комбиновати са другим доступним алатима.

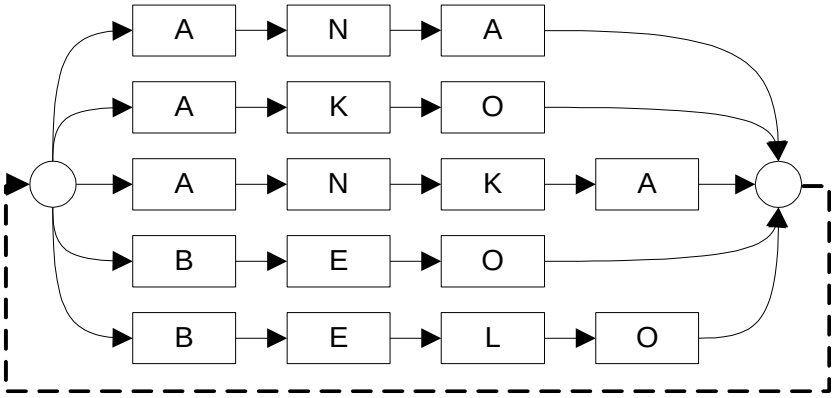
Објашњење суштине техничког решења и детаљан опис са карактеристикама, укључујући и пратеће илустрације и техничке цртеже (техничке карактеристике):

Декодер се састоји од следећих битних подмодела говора:

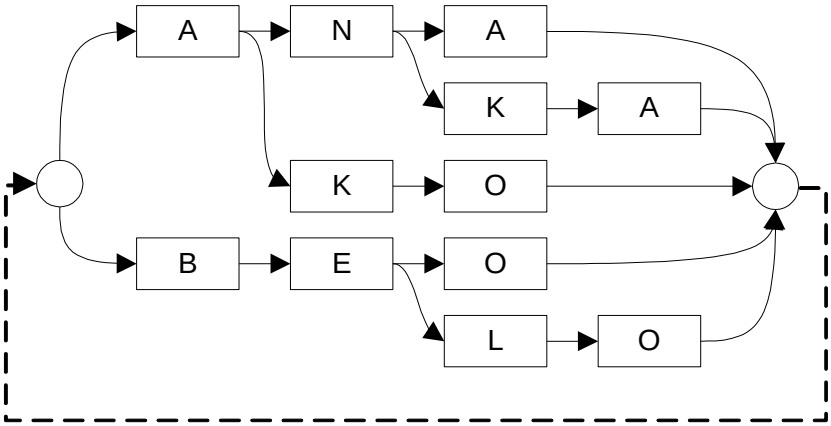
- Акустички модел – даје везу између акустичких опсервација говорног сигнала и модела алофона, који се апроксимирају моделима контекстно зависних фонема. За ове потребе користе се скривени Марковљеви модели (*hidden Markov models*) у комбинацији са мешавинама Гаусових расподела вероватноћа.
- Модел начина изговора – дефинише пресликавање речи у секвенце акустичких модела.

- Језички модел – дефинише могуће секвенце речи, односно њихове вероватноће појављивања у говору, специфичне за задати корпус. Ово је моделовано преко *N*-грама (униграма, биграма и триграма).

Када је реч о великим речницима, сваки од поменутих подмодела је сам за себе такође прилично велике сложености, те је за препознавање у реалном времену неопходно обезбедити ефикасну структуру која ће смањити простор претраге. У случају малих речника, свака реч се представља као линеарна секвенца фонема, независно од других речи (слика 1). Овакав тзв. линеарни лексикон је непрактичан за велике речнике јер непотребно повећава простор претраге, пошто ће многе подсеквенце фонема бити заједничке за велики број речи. Дељење произвољних подсеквенци фонема би значајно закомпликовало алгоритам претраге (идентификацију речи које чине неку путању), тако да се користи решење у којем речи деле сегменте са почетка речи. Оваква организација се назива лексичко или префиксно стабло (слика 2).



Слика 1. Мрежа речи (линеарни лексикон)



Слика 2. Лексичко (префиксно) стабло

Очигледна је редукција простора претраге. Са друге стране, код лексичког стабла, за разлику од линеарног лексикона, информацију о речи кроз коју путања пролази могуће је добити тек када путања уђе у последњи фонем тренутне речи, чиме се број хипотеза иницијално повећава због немогућности правовремене примене модела језика. Овај недостатак се превазилази применом тзв. факторисаног модела језика.

У случају када је примењивани модел језика сложенији од униграма, вероватноћа појаве неке речи зависи од претходних речи. Тада се вероватноће модела језика не могу директно повезати са одговарајућим чворовима у лексичком стаблу. Реализовано решење узима у обзир да стабло дефинише само могуће путање, а факторисане вероватноће модела језика се одређују на основу података у ком чвору лексичког стабла се активна хипотеза (потенцијална путања) налази и кроз које речи је претходно прошла.

Структура лексичког стабла је реализована кроз класу *CLexicalTree*. Основни податак члан класе је *m_PronemeNodes*, који садржи све чворове стабла. Информације о структури стабла се чувају локално у појединачним чворовима, који су реализовани као класа *CLexicalTreeNode*. Сваки чвор садржи следеће информације – ком фонему одговара чвор, којим могућим речима одговара чвор, који су следећи чворови и који су одговарајући акустички модели за сваки од чворова следбеника. Класа стабла осим самих чворова чува и информације о томе који су чворови на почетку речи (стартни), који су чворови на крају речи (завршни) и који чворови одговарају тишини. Тишине имају посебан третман приликом декодовања. Постоје три чвора који представљају тишину – тишина на почетку реченице, тишина на крају реченице и пауза између речи. Прва наведена нема претходне чворове, друга нема следеће чворове, док тишина између речи није обухваћена моделом језика и стога је треба игнорисати при израчунавању вероватноћа модела језика.

Лексичко стабло се генерише на основу речника изговора, који се претходно учитава из улазне текстуалне датотеке у структуру која је дефинисана класом *CWordDictionary*. Она садржи речи у њиховом писаном и изговорном облику. Под писаним обликом подразумева се облик речи који се јавља у тексту, док изговорни облик представља низ одговарајућих фонема. Пошто српски језик има фонетско писмо, ова два облика се у великој мери поклапају, а евентуалне разлике су углавном последице увођења нових гласова ради бољег акустичког моделовања (нпр. разликовање наглашених од ненаглашених вокала, увођење неутралног вокала „шва“ и слично). Једна реч може имати више изговорних облика. На основу њих се формира лексичко стабло, тако да иста реч може образовати више путања у стаблу. Чворови за тишине се накнадно додају у стабло због својих посебности.

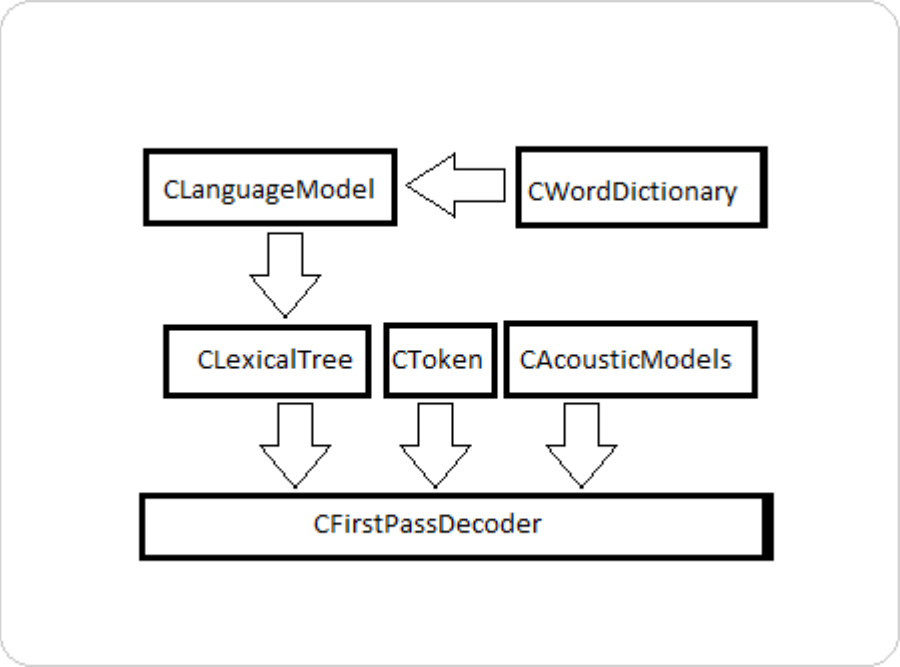
Вероватноће појаве речи из речника изговора чувају се у класи *CLanguageModel*. Тренутно је подржано моделовање униграма, биграма и триграма, али се проширивањем функција за учитавање модела језика постојећи код може лако прилагодити и за коришћење сложенијих *N*-грама. Модел језика се задаје улазном текстуалном датотеком у стандардном ARPA формату, у ком се поред условне вероватноће *N*-грама наводе и условне вероватноће једноставнијих модела и одговарајући *backoff* коефицијенти, који се користе за израчунавање условних вероватноћа које нису експлицитно задате. У циљу убрзања рада декодера, свака реч има јединствени идентификатор (број) са којим се интерно оперише у оквиру програма (уместо са низовима карактера). Конверзија идентификатора у реч се врши у последњем кораку декодовања, при генерисању излаза за корисника.

Овај декодер се заснива на тзв. *token-passing* алгоритму, који је заправо модификована варијанта Витербијевог алгоритма. Он спада у алгоритме класе „прво у ширину“ (*breadth-first*), који истовремено проширују више путања (а не само у датом тренутку највероватнију). У случају великих речника, неопходно је број потенцијалних путања држати под контролом. У овој варијанти алгоритма, информације о путањи, претходном чвору у трелису и акумулираној изгледности памте се, уместо на нивоу чвора трелиса, на нивоу хипотезе о путањи, која се пак чува у оквиру токена.

Токен је реализован преко класе *CToken*, која садржи следеће податке – акумулирану изгледност, претходни токен, тренутак (редни број опсервације) на који се односи изгледност, припадајућу реч, чвор лексичког стабла у ком се налази, наредни чвор у лексичком стаблу, одговарајући акустички модел и акустичко стање у оквиру њега, идентификатор путање и идентификатор валидности токена. У реализованом алгоритму се, дакле, не памти информација о претходном стању, већ о чвору лексичког стабла, односно речи. Пошто је сваког тренутка позната путања свих хипотеза, много је лакше реализовати препознавач за неколико најбољих хипотеза (*N-best*) у односу на оригинални Витербијев алгоритам. Информације о позицијама у лексичком стаблу и акустичком моделу неопходне су за пропагацију токена кроз стабло (у наредно стање модела или наредни чвор стабла). Када се обраде све опсервације, као потенцијалне хипотезе се узимају оне које се тада налазе у крајњим стањима модела крајњих чворова у стаблу. Спречавање експанзије броја путања врши се спајањем оних које у датом тренутку заврше у истом стању истог чвора, а имају и исту секвенцу претходећих речи (оставља се она са највећом изгледношћу). Поменута секвенца је садржана у идентификатору путање.

Пропагација токена и само декодовање реализовано је у оквиру класе *CFirstPassDecoder*. Ова класа осим листе текућих токена и листе претходних токена (историје) обједињује акустичке моделе са лингвистичким моделима (лексичко стабло и модел језика).

Учитани акустички модели су представљени класом *CAcousticModels*. Подржани су модели засновани на скривеним Марковљевим моделима и мешавинама Гаусових расподела вероватноћа (било са дијагоналним или пуним коваријансним матрицама). Ради универзалности декодера, улазни формат осим АлфаНум-овог посебног формата може бити и HTK текстуални, као и HTK бинарни формат. Улазни скуп модела се редукује на потребни скуп који је дефинисан речима у речнику изговора ради убрзања рада и оптималнијег коришћења меморијских ресурса. Комуникација са класом декодера врши се преко њеног податка члана *m_vAcousticProb* који садржи вектор вредности акустичких изгледности за текући тренутак и функције за враћање вероватноћа преласка у наредно стање, односно останка у текућем акустичком стању (претпостављена је секвенцијална структура модела). Слика 3 приказује дијаграм повезаности свих класа који се користе у оквиру декодера.



Слика 3. Дијаграм класа у оквиру декодера

Већ је споменуто да је реализација декодера за велике речнике немогућа у реалном времену без одбацивања мање вероватних хипотеза (тзв. *pruning*). Ово се ради на два начина – на основу акумулиране изгледности токена (одбацивање свих путања са изгледношћу која је за више од дате фиксне вредности мања од изгледности највероватније путање) и на основу максималног броја активних токена. Други начин захтева сортирање токена. За брзу реализацију тога, користи се још један податак члан класе декодера који садржи показиваче на токене из листе текућих токена, који се сортирају уместо оригиналних токена. Критична места су места где долази до гранања – чворови лексичког стабла, а посебно завршни чворови), па се морају посебно

третирати. То се ради увођењем пенала за које се акумулирана изгледност умањује при преласку токена у нови чвор, односно у нову реч. Посебне вредности ових пенала имају чворови тишине. Сви наведени параметри се морају пажљиво одабрати, да не би довели до грешака у декодовању.

Декодовање почиње иницијализацијом, при којој се заузму потребни меморијски ресурси, повезују модели и поставља први токен у чвор стартне тишине. Потом се за сваку опсервацију из улазне секвенце примењује следећи низ корака – пропагација активних токена, сажимање (спајање) сличних токена ради спречавања експанзије броја путања, евалуација акустичких изгледности и *pruning*. Начин пропагације токена зависи од акустичког стања у ком се токен налази. Ако се не налази у крајњем стању одговарајућег акустичког модела, пропагација се своди на клонирање датог токена и корекцију изгледности и тренутка ова два токена зависно од вероватноћа останка/прелаза и текућег времена. Токен који ће прећи у наредно стање такође повећа и вредност редног броја акустичког стања за један. Ако се пак налази у крајњем стању, неопходно је вишеструко клонирање токена за сваки од наредних чворова лексичког стабла. Врши се и ажурирање изгледности токена који остаје у текућем стању за вероватноћу останка у стању, док се за остале токене изгледност увећава за вредности логаритма одговарајућих факторисаних вероватноћа модела језика. За њих се такође постављају и нове вредности за тренутни чвор, следећи чвор, акустички модел и редни број стања, које се добијају преко информација у одговарајућем чвору лексичког стабла. Ако токен улази у нову реч, поставља му се и нова одговарајућа реч и идентификатор путање, а сам токен се смешта у историју (листу претходних). Након пропагације врши се сажимање токена у истом акустичком стању истог чвора са истом предсеквенцом речи у један токен, чиме се омогућава даља пропагација само токена са највећом изгледношћу. Евалуација акустичких изгледности подразумева детекцију активних стања на основу листе активних токена и израчунавање акустичке изгледности за та стања преко одговарајуће функције класе *CAcousticModels*. Затим се још врши корекција изгледности активних токена на основу добијених резултата и на крају циклуса следи *pruning* на претходно описане начине.

Након обраде свих опсервација, од токена који се налазе у завршном чвору тишине за крај реченице, бира се онај са највећом изгледношћу (или неколико њих у случају *N-best*), па се преко историје токена коришћењем поља за претходни токен врши пропагација уназад док се не стигне до тишине за почетак реченице. На основу припадајућих речи, тренутака и изгледности одговарајућих токена у добијеном низу могуће је реконструисати секвенцу препознатих речи, границе између речи и изгледност појединих речи, и те резултате приказати кориснику.

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

У оквиру овог декодера за препознавање континуалног говора из великих речника имплементиране су технике за ефикасно препознавање природног, континуалног говора на речницима и до неколико десетина хиљада речи. Он се може применити у многим системима, као што су системи за диктирање, системи за транскрипцију аудио записа у одговарајући текст, позивни центри и дијалошки системи. Систем је у употреби у неколико предузећа на територији Републике Србије, а у плану је интеграција овог система у још неколико предузећа у земљи и окружењу.

Корисницима су доступне следеће могућности:

- *Препознавање континуалног говора у реалном времену на речницима до 60000 речи;*
- *За системе зависне од говорника, на говорним сигналима микрофонског квалитета, са речником од око 10000 речи, тачност препознавања на нивоу речи износи око 96%;*
- *Може да функционише и у систему независном од говорника, где на говорним сигналима телефонског квалитета, на малом скупу речи (око 250) и са граматиком независном од контекста, тачност препознавања на нивоу речи је око 98%.*

Докази (прилози):

- Писано мишљење два рецензента, експерта из области техничког решења.
- Потврде предузећа АлфаНум и Аксон – интелигентни системи.
- 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 2012. године.

Подносилац пријаве:

Проф. др Владо Делић
Руководилац пројекта ТР32035

ALFANUM
GOVORNE TEHNOLOGIJE

AlfaNum d.o.o., Novi Sad
Trg Dositeja Obradovića 6
021/4750204
PIB: 102948614
Mat.br. 08790949
Tekući račun: 340-11002697-66
Registarski broj: 002890219
Šifra delatnosti: 6201

POTVRDA

Ovim potvrđujemo da je septembra 2012. godine u preduzeću AlfaNum d.o.o. pušten u rad dekoder za prepoznavanje kontinualnog govora na velikim rečnicima koji smo sami razvili i koristi se za potrebe internog testiranja i daljeg razvoja rešenja. Rešenje je inicijalno instalirano za potrebe testiranja, ali ga preduzeće AlfaNum koristi i za sopstvene potrebe.

U Novom Sadu
18. januar 2013

Darko Pekar, dipl.ing. , Direktor

axon inteligentni sistemi

Kozaračka 17, 21000 Novi Sad, Srbija
Telefon: +381 21 475 0204

Web: www.axon-is.co.rs
E-mail: office@axon-is.co.rs

POTVRDA

Ovim potvrđujemo da je septembra 2012. godine u preduzeću Axon – inteligentni sistemi d.o.o. pušten u rad dekoder za prepoznavanje kontinualnog govora na velikim rečnicima koji je razvila kompanija AlfaNum d.o.o. i koristi se za potrebe internog testiranja i razvoja naših sopstvenih rešenja za mobilne platforme. Rešenje je inicijalno instalirano za potrebe testiranja i razvoja, ali ga preduzeće Axon koristi i za sopstvene potrebe.

U Novom Sadu
18. januar 2013

Vesna Rehner , Direktor

axon inteligentni sistemi doo novi sad

РЕЦЕНЗИЈА ТЕХНИЧКОГ РЕШЕЊА

Подаци о техничком решењу:

Назив техничког решења:	Декодер за препознавање континуалног говора на великим речницима
Аутори техничког решења:	Никша Јаковљевић, Драгиша Мишковић, Едвин Пакоци, Роберт Мак, Бранислав Поповић
Реализатори:	Факултет техничких наука и АлфаНум у Новом Саду
Пројекти на којим је развијено:	„Развој дијалошких система за српски и друге јужнословенске језике“, ТР32035 код МПНТР (2011-2014)
Област на коју се односи:	Електроника, телекомуникације и информационе технологије
Корисници:	„АлфаНум“ и „Аксон – интелигентни системи“, Нови Сад
Категорија техничког решења:	Ново техничко решење – прототип (М85)

Подаци о рецензенту:

Име, презиме и звање:	Др Снежана Гудурић, редовни професор
Ужа научна област за коју је изабран у звање, датум избора у звање и назив факултета:	изабрана у звање редовног професора 21.07.2009. год. на Филозофском факултету у Новом Саду, за у.н.о. Филолошке науке
Установа где је запослен:	Филозофски факултет Универзитета у Новом Саду

Стручно мишљење рецензента:

Резултат научно-истраживачког рада „**Прототип декодера за препознавање континуалног говора на великим речницима**“ испуњава услове за признање својства техничког решења и то као **новог техничког решења - прототипа (М85)**, све у смислу одредби које се односе на техничка решења у Правилнику који је 21.03.2008. године донео Национални савет за научни и технолошки развој Републике Србије.

Образложење за техничко решење (ТР):

- ТР представља основу за системе за препознавање континуалног говора који имају многоструку примену - у системима за диктирање и транскрипцију аудио снимака говора, у позивним центрима, дијалошким системима.
- ТР омогућава препознавање континуалног говора у реалном времену на речницима и до 60000 речи, за системе зависне од говорника тачност износи у просеку и до 96% речи, док одличне резултате постиже и за мање речнике у системима независним од говорника.
- ТР је развијено коришћењем научних метода на пројектима МПНТР, претежно на сопственој опреми и без коришћења туђе патентне/лиценцне документације. Као резултат рада на овом ТР произашло је и неколико научних радова презентованих на конференцијама.
- ТР се већ примењује код горе наведених корисника у Србији, и представља прво решење проблема са оваквим скупом карактеристика.

У Новом Саду, 10..01.2013. године.

Др Снежана Гудурић

РЕЦЕНЗИЈА ТЕХНИЧКОГ РЕШЕЊА

Подаци о техничком решењу:

Назив техничког решења:	Декодер за препознавање континуалног говора на великим речницима
Аутори техничког решења:	Никша Јаковљевић, Драгиша Мишковић, Едвин Пакоци, Роберт Мак, Бранислав Поповић
Реализатори:	Факултет техничких наука и АлфаНум у Новом Саду
Пројекти на којим је развијено:	„Развој дијалошких система за српски и друге јужнословенске језике“, ТР32035 код МПНТР (2011-2014)
Област на коју се односи:	Електроника, телекомуникације и информационе технологије
Корисници:	„АлфаНум“ и „Аксон – интелигентни системи“, Нови Сад
Категорија техничког решења:	Ново техничко решење – прототип (М85)

Подаци о рецензенту:

Име, презиме и звање:	Проф. др Слободан Јовичић
Ужа научна област за коју је изабран у звање, датум избора у звање и назив факултета:	изабран у звање редовног професора 01.07.2011. године на Електротехничком факултету у Београду, за у.н.о. Говорна комуникација
Установа где је запослен:	Електротехнички факултет у Београду

Стручно мишљење рецензента:

Резултат научно-истраживачког рада „**Прототип декодера за препознавање континуалног говора на великим речницима**“ испуњава услове за признање својства техничког решења и то као **новог техничког решења - прототипа (М85)**, све у смислу одредби које се односе на техничка решења у Правилнику који је 21.03.2008. године донео Национални савет за научни и технолошки развој Републике Србије.

Образложење за техничко решење (ТР):

- ТР представља декодер за препознавање статистички моделованог говора на широком скупу речи, модуларне структуре и добро документован. Основа је за бројне системе који користе препознаваче говора.
- За сада се примењује тако што се позива из командне линије оперативног система. Може у реалном времену да препознаје говор са изузетном тачношћу на скупу од неколико десетина хиљада речи, зависно и независно од говорника.
- ТР је развијено коришћењем научних метода на пројектима МПНТР, претежно на сопственој опреми и без коришћења туђе патентне/лиценцне документације. У току развоја овог ТР произашло је и неколико научних радова презентованих на конференцијама.
- ТР се већ примењује код горе наведених корисника у Србији, и представља прво решење проблема са оваквим скупом карактеристика.

У Београду, __.01.2013. године.

др Слободан Јовичић

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: 2013-01-30

ИЗВОД ИЗ ЗАПИСНИКА

Наставно-научног већа Факултета техничких наука у Новом Саду, на 4. редовној седници одржаној дана 30.01.2013. године, донело је следећу одлуку:

-непотребно изостављено-

Тачка 15.1.35.: Питања научноистраживачког рада и међународне сарадње / верификација нових техничких решења

Одлука

На основу позитивног извештаја рецензената верификује се ***техничко решење (М85) под називом:***

ДЕКОДЕР ЗА ПРЕПОЗНАВАЊЕ КОНТИНУАЛНОГ ГОВОРА НА ВЕЛИКИМ РЕЧНИЦИМА.

Аутори техничког решења: Никша Јаковљевић, Драгиша Мишковић, Едвин Пакоци, Роберт Мак, Бранислав Поповић.

-непотребно изостављено-

Записник водила:
Јасмина Димић, дипл. правник

Тачност података оверава:
Секретар
Иван Нешковић, дипл. правник

Декан
Проф. др Раде Дорословачки