

## Рецензия

върху дисертационен труд за придобиване на образователната и научна степен “доктор” по научна специалност: 01.01.12. Информатика в професионално направление: 4.6. Информатика и компютърни науки  
Автор на дисертационния труд: **маг. инж. Атанас Петров Узунов**

Тема на дисертационния труд:

### **ДЕТЕКЦИЯ НА ГОВОР В СИСТЕМИ ЗА РАЗПОЗНАВАНЕ НА ДИКТОРИ**

Рецензент: **проф. д-р инж. Александър Богданов Бекярски**

#### **1. Актуалност на разработвания в дисертационния труд проблем в научно и научно-приложно отношение**

Темата на представения за рецензия дисертационен труд “ДЕТЕКЦИЯ НА ГОВОР В СИСТЕМИ ЗА РАЗПОЗНАВАНЕ НА ДИКТОРИ” съвсем естествено разкрива наличието на актуалност, свързана със съществуващите и постоянно нарастващи изисквания към обработката на говорни сигнали чрез подходящи методи и подходи. С използването биометричните технологии, с цел разпознаване на диктор, чрез анализът на индивидуалните характеристики на говорещия се повишава точността на идентификация и разпознаване. Това определя актуалността и дефинира съответната цел на настоящия дисертационен труд – разработване на методи и алгоритми за определяне на робастни признаци, предназначени за алгоритми за детекция на говор и тяхното изследване в контекста на задачите за разпознаване на диктори при говорен сигнал, записан по телефонен канал. Подходът и инструментариумът, предложени при разработването на дисертацията, са свързани с детектиране на говор, чрез прилагане на комбинирани методи и алгоритми за обработка и анализ на говорни сигнали и използване на априорни знания. Тази тенденция е залегнала в основата на изразената в темата на дисертацията актуалност и перспективност в научно и научно-приложно отношение. В съответствие с тази оценка за актуалност следва да се подчертае и съответната насоченост в дисертацията за предлагане, изследване и практическо използване на нови методи в обработката на говорни сигнали.

#### **2. Степен на познаване състоянието на проблема и творческа интерпретация на литературния материал**

Същността на Глава 1 на дисертационния труд директно разкрива, обосновава и доказва отличното познаване от докторанта на състоянието на проблема – обработка на говорни сигнали, предавани по телефонен канал, с цел определяне на робустни принципи чрез анализ на индивидуалните характеристики на говорещия. Това е постигнато чрез съответна подходяща

творческа интерпретация на литературния материал, задълбочено анализиране на настоящото ниво на научните изследвания, разработки, експерименти и практически реализации в областта на обработката на говорни сигнали. В процеса на извършване на обстоен преглед и критичен анализ на съществуващите методи и изследвания в областта на обработката на говорни сигнали са използвани общо **151** литературни източника, от които **149** са на латиница и **2** на кирилица. Всички те са актуални в настоящия момент научни публикации, включени в приложения списък на литературните източници към дисертацията и са от реномирани списания, издания от международни конференции и електронни източници от водещи, изтъкнати в настоящия момент автори, чужди учени и специалисти в областта на обработката на говорни сигнали.

### **3. Съответствие на избраната методика на изследване и поставената цел и задачи на дисертационния труд с постигнатите приноси**

Забелязва се пълно и добре обосновано съответствие на избраната методика на изследване и поставената цел и задачи на дисертационния труд с постигнатите приноси. В това твърдение следва да се включи обосноваването на целите и задачите в дисертационния труд, напълно основателно дефинираната значимост и необходимост в научно и научно-приложно отношение от изследвания в областта на обработката на говорни сигнали. Предложен е правилен подход, по който докторантът е формулирал точно и в синтезиран вид както целта, така и задачите в своя дисертационен труд, както следва:

**Цел на дисертацията:** Целта на дисертационния труд е формиране на робастни признаци, предназначени за алгоритми за детекция на говор и тяхното изследване в контекста на задачите за разпознаване на диктори от говорен сигнал, записан по телефонен канал.

В съответствие с дефинираната цел на дисертационния труд са формулирани следните задачи на дисертацията:

1. Дефиниране на робастни признаци, предназначени за детекция на говор и базиращи се на свойствата на спектралната автокорелационна функция и спектъра на групово закъснение и изследване на техните характеристики
2. Разработване на подход за определяне на граничните точки на говорно съобщение включващ алгоритъм за изчисляване на адаптивни прагови стойности и детерминиран краен автомат
3. Разработване на алгоритми за определяне на гранични точки и експериментално изследване на тяхната ефективност с предложените в дисертацията признаци при верификация на диктори с фиксирани фрази
4. Разработване на алгоритми за детекция на говорни сегменти и експериментално изследване на тяхната ефективност с предложените в дисертацията признаци при независима от текста идентификация на диктори

### **4. Кратка аналитична характеристика на естеството и оценка на достоверността на материала, върху който се градят приносите на дисертационния труд**

Дисертационният труд е с обем **160** страници. включва увод, **пет** глави, тясно свързани с формулираните основни задачи, заключение (резюме на получените резултати), списък на основните приноси, списък на публикациите по дисертацията, използвана литература, списък на използваните съкращения. Цитирани са общо **151** литературни източници, като **149** от тях са на латиница и **2** на кирилица. В дисертацията има общо **48** фигури и **27** таблици. Номерата на фигурите и таблиците в автореферата съответстват на тези в дисертационния труд. Реализирането на правилно поставената цел и съответни задачи на дисертационния труд е обект на следващите пет глави на дисертацията, между които е постигната последователност в изложението, логическа връзка и методичност в представяне на новите важни, относно приносите на дисертацията, теоретични и практически решения в областта на темата на дисертацията.

В Глава 2 “Дефиниране на признаци за детекция на говор, използващи свойствата на САКФ и СГЗ” са разгледани някои характеристики на спектралната автокорелационна функция (САКФ), получена чрез FFT-спектъра. Предложен е метод, при който чрез прилагане на делта-филтър върху спектралната автокорелационна функция (САКФ) е получена делта спектрална автокорелационна функция. Анализирано е изменението на GDS при зашумени с адитивен шум говорни сигнали - на базата само на свойствата на делта спектралната автокорелационна функция и чрез комбинирането ѝ с модифицирания спектър на групово закъснение. Въз основа на направения анализ са предложени пет признака за детекция на говор - MD, log-GDMD, lin-GDMD, BMD и MMD. Първите три са предназначени за детекция чрез анализ на времеви контури, а последните два - за детекция чрез алгоритми за разпознаване.

Въз основа на проведените в Глава 2 изследвания и съответен анализ в Глава 3 “Алгоритми за определяне на гранични точки при зависима от текста верификация на диктори. Експериментално изследване” е извършен сравнителен експериментален анализ на ефективността на предложените в гл. 2 признаци предназначени за детекция на говор чрез анализ на времеви контури. За референтни признаци са избрани: признак, получен чрез комбинация на енергията на сигнала и спектралната ентропия (Energy-Entropy (EE) parameter); спектрална ентропия с нормализиран спектър (Spectral Entropy with Normalized frame Spectrum – SENS parameter); модифицирана енергия на Teager (Modified Teager’s Energy – MTE parameter) и дълговременна спектрална дивергенция (Long-term Spectral Divergence -LTSD parameter). Предложени, разработени и изследвани са методи и алгоритми за определяне на граничните точки на кратки фрази.

На базата на получените експериментални резултати са направени следните три заключения: Първо – детекторите на базата на log-GDMD признака във всички тестове превъзхождат тези на базата на LTSD; Второ – точността на детекция при използване на краен автомат с адаптивни прагове винаги превъзхожда, при един и същи признак, тази получена чрез hangover алгоритъма и Трето – от гледна точка на грешката при 24 верификации в повечето случаи

детектора с краен автомат и адаптивни прагове превъзхожда, при един и същ признак, този с hangover алгоритъма, но разликата между тях не статистически значима. Експериментално е изследвана точността на детекция чрез анализ на разликите между ръчно определените гранични точки и тези, получени от предложените алгоритми. Установено е как различните алгоритми за детекция влияят върху точността на разпознаване при зависима от текста верификация на диктори.

Предложените и разработени в Глава 3 методи и алгоритми са основните, чрез които е теоретично обосновано и аргументирано постигането на целта в дисертационния труд - алгоритми за детекция на говор и тяхното изследване с цел идентификация и разпознаване на диктори.

В Глава 4 „Алгоритми за детекция на говор при независима от текста идентификация на диктори. Експериментално изследване” е извършен сравнителен експериментален анализ на ефективността на предложените в гл. 2 признаци, предназначени за детекция на говор. Експериментално са изследвани два алгоритма за детекция на говор –VAD-1 и VAD-2. Изследвани са два детектора на говор, които в действителност са бинарни класификатори. В алгоритъма за идентификация на диктори, независимо от текста като класификатор, се използва многослоен перцептрон с един скрит слой и алгоритъм за обучение чрез обратно разпространение на грешката. Експериментално определените резултати от алгоритмите за детекция на говор и идентификация на диктори са представени в табличен и графичен вид. Направени са изводи за тяхната ефективност. Изследвано е влиянието на алгоритмите за детекция на говор върху точността на разпознаване в система за независима от текста идентификация на диктори. Експериментите са реализирани с говорни данни на български език, записани по телефонен канал.

В глава 5 “BG-SRDat – корпус с говорни данни, записани по телефонен канал, и предназначен за разпознаване на диктори” е описан корпуса BG-SRDat (Bulgarian language Speaker Recognition DATa), съдържащ говор, записан по телефонен канал (стационарни и мобилни телефони и чрез VoIP) и включващ фрази и разговори на български и само фрази на английски език. Той е използван за изследвания в областта на верификация на диктори чрез фиксирани фрази (на български и английски), независимо от текста идентификация на диктори и детекция на говор. Основна тенденция при бъдещо развитие на корпуса ще бъде постепенното му преобразуване в корпус, съдържащ говорни данни, получени само от мобилни устройства.

## **5. Научни и/или научно-приложни приноси на дисертационния труд**

Основните резултати, постигнати в настоящия дисертационен труд са обобщени и представени в синтезиран вид под формата на следните научно-приложни и приложни приноси:

### **Научни приноси:**

1. Предложен е метод, при който чрез прилагане на делта-филтър върху спектралната автокорелационна функция е получена т.н. делта спектрална автокорелационна функция (Глава 2, т. 2.1.3)

2. Извършен е теоретичен анализ на изменението на спектъра на групово закъснение при зашумени с адитивен шум говорни сигнали (Глава 2, т. 2.2.3).
3. Предложени са следните подходи: за изчисляване на признаци за детекция на говор базиращ се на свойствата на делта спектралната автокорелационна функция. Чрез този подход са дефинирани три признака (Глава 2, т. 2.1.4); за изчисляване на признаци за детекция на говор базиращ се комбинация на модифицирания спектър на групово закъснение и делта спектралната автокорелационна функция (Глава 2, т. 2.2.4); за определяне на гранични точки на говорно съобщение, включващ алгоритъм за изчисляване на адаптивни прагови стойности и детерминиран краен автомат (Глава 3, т. 3.4.2-3).

#### **Научно-приложни приноси:**

1. Извършен е сравнителен експериментален анализ на: предложените в гл. 2 признаци спрямо избрани референтни такива, при което е предложено използване на Евклидово разстояние между  $Z$ -нормализирани времеви контури, изчислени за всеки признак, съответно от чист и от зашумен сигнал (Глава 3, т. 3.3); на ефективността на предложените в гл. 2 признаци спрямо избрани референтни такива при използването им в разработените алгоритми за определяне на гранични точки (Глава 3, т. 3.6.3 и Глава 3, т.3.6.4); на ефективността на предложените в гл. 2 параметри спрямо избрани референтни такива при използването им във VAD-1 и VAD-2 (Глава 4, т.4.6.5 и т.4.8.4, Глава 4, т.4.7 и т.4.9).
2. Разработени са следните алгоритми: три алгоритъма за определяне на гранични точки, базиращи се на предложения подход и формирани съобразно използваните времеви контури (Глава 3, т. 3.5); два алгоритъма за детекция на говорни сегменти с използване на класификатор на базата на многослойна невронна мрежа скаларни, векторни признаци и прагова логика (Глава 4, т.4.6 и т.4.8).

#### **Приложни приноси:**

1. Разработване, с подчертано участие на докторанта, като главен научен консултант във фирмата Fadata, Ltd., на проект “Програмен интерфейс Speaker Recognition API (SR-API)”, който е използван (в Глава 3) в програмната реализация на алгоритъма за определяне на гранични точки и в две приложения за верификация на диктори – DTW и HMM.

### **6. Оценка на степента на личното участие на докторантката в приносите**

Личното участие на докторантът в представените научно-приложни и приложни приноси към дисертацията е неоспоримо и потвърдено, както от изложението в дисертацията и автореферата, така и от публикационната дейност (6 научни публикации), където докторантът няма съавтори.

### **7. Преценка на публикациите по дисертационния труд**

Основните постижения в дисертационния труд са популяризирани в 6 научни публикации, от които една е реферирана в **Web of Science**, три са реферирани в **Scopus**. Всички публикации са с теоретично и приложно значение,

свързани са с дисертацията и с професионалното направление 4.6. Информатика и компютърни науки. Докторантът е представил справка за забелязани цитирания – общо 25 броя. Приемам публикационната дейност като напълно достатъчна по обем, на високо научно ниво и популяризирана в достатъчна степен в национален и международен научен план.

#### **8. Използване на резултатите от дисертационния труд в научната и социалната практика**

Като реално доказателство за използването на резултатите от дисертационния труд в научната и социалната практика следва да се приеме декларираното участие на докторантът в представената „Справка за участие в договори, свързани с темата на дисертацията“: докторантът е участник в два договора с Фонд Научни изследвания „Биометрични параметри за идентификация” и „Методи и алгоритми за анализ на комбинирана биометрична информация”; Договор от VI рамкова програма - №507634 (2004-2007), BIOSECURE – Network of Excellence - проект от 6-та рамкова програма за идентификация по биометрични показатели; два вътрешни договора на ИИТ-БАН «Идентификация на биометрични параметри» и «Разпознаване на биометрични параметри». Докторантът е бил главен научен консултант в договор за разработка на приложен интерфейс за разпознаване на диктори - Speaker Recognition API (SR-API). Участието на докторанта в тези договори, тясно свързани с темата на дисертационния труд, както и успешната им реализация, ясно показват стремежа на докторантът към реално прилагане на разработените в дисертацията алгоритми и програмни приложения за детектиране на говор в системи за разпознаване.

#### **9. Оценка на съответствието на автореферата с изискванията за изготвянето му, както и на адекватността на отразяване на основните положения и приносите на дисертационния труд**

Прецизният преглед на автореферата към дисертационния труд показва наличие на пълно съответствието на автореферата с изискванията за изготвянето му, както и на адекватността на отразяване на основните положения и приносите на дисертационния труд. Авторефератът към дисертационния труд напълно основателно може да се оцени и характеризира като синтезиран вариант на дисертацията, отразяващ вярно нейното основно съдържание, постигнатите теоретични и практически резултати и дефинираните научно-приложни и приложни приноси, съдържащи се в пълния текст на дисертационния труд.

#### **10. Мнения, препоръки и бележки**

Всички теоретични и практически постановки в дисертационния труд са представени аргументирано, правилно и подходяща в методично отношение последователност. Затова при рецензирането на настоящия дисертационен труд могат да се посочат само незначителни редакционни неточности и пропуски, някои от които са следните:

1. В изводите към първа глава (1.4. Заключение) би следвало в критичен аспект да се подчертае на базата на обширния преглед на съществуващите методи необходимостта от нови изследвания в областта на методите и алгоритмите за детектиране на говор, за да се аргументират по-ясно поставената цел и задачи на дисертацията.

2. Представените няколко научно-приложни приноси (№ 1, № 3, № 4, № 6 и № 7), свързани с провеждане на "сравнителен експериментален анализ", могат да се обединят (без това да принижават тяхната значимост), тъй като всички те се отнасят до "предложените в гл. 2 признаци спрямо избрани референтни такива", но чрез прилагане на различни критерии за сравнение.

3. Научните и научно-приложни приноси следва да бъдат представени в сравнителен аспект спрямо съществуващи подобни методи, подходи и алгоритми, за да се подчертае още по-ясно конкретния принос и порожителен ефект от разработените в дисертацията нови или модифицирани методи, подходи и алгоритми.

4. Считаю, че разработените в дисертацията методи, подходи и алгоритми имат стойност и практическо приложение при детектиране на говор за разпознаване на диктори не само от говорни сигнали, записани по телефонен канал, което е полезно и би следвало да бъде отразено в приносите към дисертацията.

Предполагам, че тези забележки и препоръки ще бъдат ценни и полезни в бъдеща научна и изследователска дейност на докторанта.

## II. Заключение

Считаю, че общата положителна оценка на теоретичните и практическите резултати, постигнати при разработката на настоящия дисертационен труд, дефинирани като научно-приложни и приложни приноси и отразени в достатъчен брой научни публикации и в подходящи научни списания и конференции, са напълно достатъчно основание за ясно положително заключение относно квалификацията на докторантът и неговите потвърдени в дисертационния труд качества на учен в избраната от него научна област. Затова предлагам на уважаемото научно жури да присъди образователната и научна степен "доктор" на маг. инж. Атанас Петров Узунов в професионално направление 4.6. Информатика и компютърни науки.

Дата: 11.06.2020 г.

ЧЛЕН НА ЖУРИТО: .....

проф. д-р Александър Богданов Бекареки

**NOT FOR  
PUBLIC RELEASE**