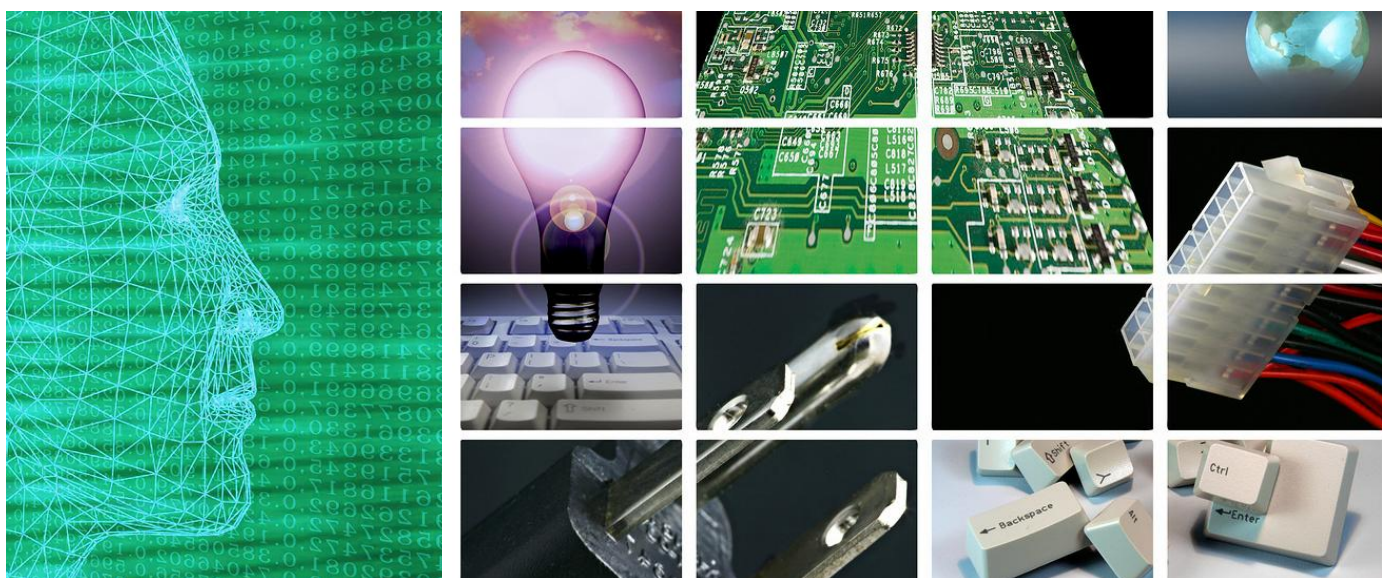




ТЕХНИЧЕСКИ УНИВЕРСИТЕТ - ВАРНА  
TECHNICAL UNIVERSITY OF VARNA

Година XVI, Брой 1/2018

# КОМПЮТЪРНИ НАУКИ И ТЕХНОЛОГИИ



Bulgaria Communications Chapter

Faculty of Computing & Automation

## COMPUTER SCIENCE AND TECHNOLOGIES

*Пета научна конференция с международно участие  
"Компютърни науки и технологии"  
28 - 29 септември, 2018 г., Варна, България*

Year XVI, Number 1/2018

# Компютърни науки и технологии

## Издание

на Факултета по изчислителна техника и  
автоматизация  
Технически университет - Варна

**Редактор:** доц. д-р Ю. Петкова  
**Гл. редактор:** доц. д-р Н. Николов

## Редакционна колегия:

проф. д.н. Л. Личев (Острава)  
проф. д.н. М. Илиев (Русе)  
проф. д-р М. Лазарова (София)  
проф. д-р Г. Спасов (Пловдив)  
доц. д-р П. Антонов (Варна)  
доц. д-р Р. Вробел (Вроцлав)  
доц. д-р Т. Ганчев (Варна)  
доц. д-р Н. Атанасов (Варна)  
доц. д-р М. Маринов, (Русе)

**Печат:** ТУ-Варна

## За контакти:

Технически университет - Варна  
ФИТА  
ул. „Студентска” 1, 9010 Варна,  
България  
тел./факс: (052) 383 320  
e-mail: ned.nikolov@tu-varna.bg  
yulka.petkova@tu-varna.bg

**ISSN 1312-3335**

# Computer Science and Technologies

## Publication

of Computing and Automation Faculty  
Technical University of Varna

**Editor:** Assoc. Prof. Y. Petkova, PhD  
**Chief Editor:** Assoc. Prof. N. Nikolov, PhD

## Editorial Board:

Prof. L. Lichev, DSc (Ostrava)  
Prof. M. Iliev, DSc (Ruse)  
Prof. M. Lazarova, PhD (Sofia)  
Prof. G. Spasov, PhD (Plovdiv)  
Assoc. Prof. P. Antonov, PhD (Varna)  
Assoc. Prof. R. Wrobel (Wroclaw)  
Assoc. Prof. T. Ganchev, PhD (Varna)  
Assoc. Prof. N. Atanasov, PhD (Varna)  
Assoc. Prof. M. Marinov, PhD (Ruse)

**Printing:** TU-Varna

## For contacts:

Technical University of Varna  
Faculty of Computing and Automation  
1, Studentska Str., 9010 Varna,  
Bulgaria  
Tel/Fax: (+359) 52 383 320  
e-mail: ned.nikolov@tu-varna.bg  
yulka.petkova@tu-varna.bg

**ISSN 1312-3335**

# СПОНСОРИ НА КОНФЕРЕНЦИЯТА

 <p><a href="http://www.researchmetrics.com">http://www.researchmetrics.com</a></p>	<p><b>Рисърчметрикс</b> е международна компания с офиси в Северна Америка, Европа и Азия, в които работят над 100 инженери и специалисти. Фирмата разработва иновационни решения, работещи в областта на набора и анализа на данни и усъвършенстване на потребителското обслужване.</p>
 <p><a href="http://www.actbg.bg">http://www.actbg.bg</a></p>	<p><b>АКТ БГ ООД</b> е създадена в началото на 1997г. като 100% частна фирма с предмет на дейност: •Доставка, проектиране, инсталиране, абонаментно обслужване и сервиз на компютърно-комуникационно оборудване и офис техника; •Проектиране и внедряване на ИТ решения за високонадеждни информационни системи; •Информационна сигурност и защита на личните данни; •Изграждане на локални и глобални мрежи; •Проектиране и изграждане на системи за контрол на достъп; •Персонализиране на апаратно-програмни решения и системи.</p>
 <p><a href="https://adastracorp.com/academy">https://adastracorp.com/academy</a></p>	<p><b>Adastra Academy</b> е уникалният бранд, който отразява визията на Адастра България за успешни служители чрез непрестанно обучение в рамките на вътрешни и външни инициативи, алтернативни начини за споделяне на знания и иновации.</p>
 <p><a href="http://www.eurorisksystems.com">http://www.eurorisksystems.com</a></p>	<p><b>EuroRisk Системи ООД</b> предоставя софтуерни системи и услуги за банки, застрахователни и финансови институции и други доставчици на финансови услуги.</p>
 <p><a href="http://beehive.bg">http://beehive.bg</a></p>	<p><b>Beehive</b> е първото във Варна споделено работно място с 5 години опит в организирането на културни събития и проекти. Провежда над 100 безплатни обучения ежегодно. На фокус са мащабни и международни конференции, startup, предприемачески формати и арт проекти. Успешно реализира проекти и в други градове, сред които са TEDx, Startup Weekend, Europe Code Week. Събира и обединява идейни и креативни личности, като им предоставя среда за ефективно развитие на бизнес и творчески интереси. Като менторска програма обучава младежи в началото на своята кариера и спомага за по-лесната им бъдеща реализация.</p>
 <p><a href="http://www.bg.adastragrp.com">http://www.bg.adastragrp.com</a></p>	<p><b>АДАСТРА</b> е компания - лидер в областта на информационния мениджмънт, осигуряваща бизнес консултиране и smart data management услуги, решения и софтуер от най-висок клас.</p>
 <p><a href="http://www.taxback.com">http://www.taxback.com</a></p>	<p><b>Taxback Group</b> предлага пакет от услуги, сред които специализирани данъчни декларации, финансови и туристически консултантски услуги за физически лица и корпоративни клиенти. Компанията предлага възможности за развитие за специалисти в сферата на информационните технологии и програмирането.</p>

 <a href="http://smartpro-bg.com/">http://smartpro-bg.com/</a>	<p>Създадена преди 20 години в град Варна, <b>SmartPro Ltd.</b> е софтуерна компания, специализирана в развитието и внедряването на софтуерни решения в помощ на бизнеса. Основен на продукт е ERP системата „Интегра“ - успешно внедрена в над 200 компании.</p>
 <a href="https://www.flatrocktech.com/">https://www.flatrocktech.com/</a>	<p><b>Flat Rock Technology Ltd.</b> е софтуерна компания, развиваща се в две направления разработка на софтуерни и мобилни приложения и предоставяне на аутсорсинг услуги. Основана преди 10 години в Лондон към момента в централния офис в град Варна работят над 130 специалисти. Flat Rock Technology дава възможност на младите да започнат своя кариерен път като програмисти, маркетинг специалисти, анализатори на данни и още много други, ръководени от професионалисти в сферата.</p>
 <a href="http://www.immedis.com">http://www.immedis.com</a>	<p><b>Immedis</b> е най-бързо развиващата се международна пейрол технологична компания в света. За офиса си в град Варна Immedis набира служители на следните позиции: Net Developers, Business Intelligence and Data Analysts, Business Analysts, Quality Assurance Specialists, Wintel Server Administrators, UI/UX Design Specialists, AWS Administrators and Database Administrators.</p>



Bulgarian ACM Chapter



## *Computer Sciences and Engineering*

is the Fifth International Scientific Conference organized by the Computer Science and Engineering and Software and Internet Technologies Department, Technical University of Varna, Bulgaria in association with the Bulgarian ACM Chapter - **acmbul**. This Conference is the 32<sup>nd</sup> **acmbul** event during the last 28 years

# СБОРНИК ДОКЛАДИ

*Пета научна конференция с международно участие  
"Компютърни науки и технологии"  
28 – 29 септември 2018 г.  
Варна, България*



*Fifth International Scientific Conference  
Computer Sciences and Engineering  
28 – 29 September, 2018  
Varna, Bulgaria*

## PROCEEDINGS

<b>Председатели:</b> Христо Вълчанов Виолета Божикова	<b>Chairmen:</b> Hristo Valchanov Violeta Bozhikova
---	---

Организационен комитет	Organizing committee
<b>Председател:</b> Венета Алексиева <b>Членове:</b> Ивайло Пенев Диян Желев Николай Дуков Мая Тодорова Айдън Хъкъ Антоанета Иванова Илиян Бойчев Стефка Попова	<b>Chairman:</b> Veneta Aleksieva <b>Members:</b> Ivaylo Penev Diyan Zhelev Nikolay Dukov Maya Todorova Aydan Haka Antoaneta Ivanova Iliyan Boychev Stefka Popova

Програмен комитет	Programme committee
<b>Председател:</b> Марияна Стоева <b>Членове:</b> Валентина Кукенска Никола Николов Радослав Вробел Розалина Димова Милена Лазарова Синиша Илич Гриша Спасов Румен Трифонов Делян Генков Ирена Въллова Милко Маринов Марта Сийбауер	<b>Chairman:</b> Mariyana Stoeva <b>Members:</b> Valentina Kukenska Nikola Nikolov Radoslaw Wrobel Rozalina Dimova Milena Lazarova Sinisa Ilic Grisha Spasov Rumen Trifonov Delyan Genkov Irena Valova Milko Marinov Marta Seebauer



## СЪДЪРЖАНИЕ

## CONTENTS

	КОМПЮТЪРНИ СИСТЕМИ И МРЕЖИ		COMPUTER SYSTEMS AND NETWORKS	
1	<b>ЧУВСТВИТЕЛЕН НА ДОПИР ДВУПРЪСТОВ ИНТЕРФЕЙС ЗА МОДЕЛ НА ПРОТОТИП НА НОСИМО УСТРОЙСТВО</b> <i>Юри Димитров</i>	11	<b>TWO-FINGER TOUCH INTERFACE FOR WEARABLE DEVICES PROTOTYPE MODEL</b> <i>Yuri Dimitrov</i>	1
2	<b>СИСТЕМА ЗА СИМУЛИРАНЕ НА МЕХАНИЗМА НА ИЗГРАЖДАНЕ НА НАЙ- ПРЕДПОЧИТАНИЯ ПЪТ ПРИ 6LOWPAN ТЕХНОЛОГИЯ</b> <i>Айдън Хъкъ, Русен Василев</i>	17	<b>SYSTEM FOR SIMULATING THE CONSTRUCTION OF THE BEST ROUTE IN 6LOWPAN TECHNOLOGY</b> <i>Aydan Haka, Rusen Vasilev</i>	2
3	<b>СРАВНИТЕЛЕН АНАЛИЗ НА МЕХАНИЗМИ ЗА ПРИОРИТИЗАЦИЯ В LTE SCHEDULER</b> <i>Айдън Хъкъ</i>	24	<b>COMPARATIVE ANALYSIS OF PRIORITIZATION MECHANISMS IN LTE SCHEDULER</b> <i>Aydan Haka</i>	3
4	<b>СРАВНЕНИЕ НА ПРОИЗВОДИТЕЛНОСТТА НА ВГРАДЕНИТЕ ИНСТРУМЕНТИ ЗА ВЪВЕЖДАНЕ И ИЗВЕЖДАНЕ НА ДАННИ ОТ ПОПУЛЯРНИ БАЗИ ОТ ДАННИ</b> <i>Синиша Илич, Драгиша Милкович, Алемпийе Вельович, Ясмينا Новакович, Владимир Вельович</i>	32	<b>COMPARISON OF PERFORMANCES OF BUILT-IN TOOLS FOR LOAD AND EXPORT OF DATA TO AND FROM POPULAR DATABASES</b> <i>Sinisa Ilic, Dragisa Miljkovic, Alempije Veljovic, Jasmina Novakovic, Vladimir Veljovic</i>	4
5	<b>ДЕСКТОП ПРИЛОЖЕНИЕ ЗА ОБУЧЕНИЕ В КОДИРАНЕ НА ДАННИ И ВИЗУАЛИЗАЦИЯ НА МОДУЛАЦИИ</b> <i>Венета Алексиева, Владислав Димитров</i>	39	<b>DESKTOP APPLICATION FOR EDUCATION IN DATA ENCODING AND VIZUALIZATION OF MODULATIONS</b> <i>Veneta Aleksieva, Vladislav Dimitrov</i>	5
6	<b>ПРОЕКТИРАНЕ И РЕАЛИЗАЦИЯ НА WEB ПРИЛОЖЕНИЕ ЗА АНАЛИЗ И СРАВНЕНИЕ НА HTTP ПРОТКОЛИТЕ SPDY И HTTP/2</b> <i>Владимир Димитров, Георги Гюров</i>	46	<b>DESIGN AND IMPLEMENTATION OF WEB APPLICATION FOR ANALYSIS AND COMPARISON OF HYPERTEXT TRANSFER PROTOCOLS SPDY AND HTTP/2</b> <i>Vladimir Dimitrov, Georgi Gyurov</i>	6

7	<b>СИСТЕМА ЗА ПРОСЛЕДЯВАНЕ НА АВТОМОБИЛ ПОСРЕДСТВОМ МОБИЛНИ ТЕХНОЛОГИИ</b> <i>Георги Георгиев, Милко Маринов</i>	54	<b>VEHICLE TRACKING SYSTEM USING MOBILE TECHNOLOGIES</b> <i>Georgi Georgiev, Milko Marinov</i>	7
8	<b>NVME SSD ДИСКОВЕ И ТЯХНОТО ИЗПОЛЗВАНЕ В ОБУЧЕНИЕТО НА IT СПЕЦИАЛИСТИ</b> <i>Явор Дечев</i>	62	<b>NVME SSD DRIVES AND THEIR USE IN THE EDUCATION OF IT SPECIALISTS</b> <i>Yavor Dechev</i>	8
9	<b>ИЗПОЛЗВАНЕ НА ДЕСКТОП ВИРТУАЛИЗАЦИИ ЗА ВИРТУАЛНИ ЛАБОРАТОРИИ</b> <i>Делян Генков</i>	70	<b>USING DESKTOP VIRTUALIZATION FOR VIRTUAL LABORATORIES</b> <i>Delyan Genkov</i>	9
10	<b>УПРАВЛЕНИЕ НА ЗАХРАНВАНЕТО НА ВИРТУАЛНА ИНФРАСТРУКТУРА</b> <i>Христо Вълчанов</i>	77	<b>POWER MANAGEMENT OF A VIRTUAL INFRASTRUCTURE</b> <i>Hristo Valchanov</i>	10

	<b>СОФТУЕРНИ И ИНТЕРНЕТ ТЕХНОЛОГИИ</b>		<b>SOFTWARE AND INTERNET TECHNOLOGIES</b>	
1	<b>ЕРГНОМИЧНИ АСПЕКТИ НА ДОБИВАНЕТО НА КРИПТОВАЛУТА</b> <i>Радослав Вробел, Селестин Милуш, Пьотр Халер, Радослав Влостовски</i>	84	<b>ERGONOMIC ASPECTS OF CRYPTOCURRENCY MINING</b> <i>Radosław Wróbel, Celestyn Miluch, Piotr Haller, Radosław Włostowski</i>	1
2	<b>КЪМ МОДУЛНА И АДАПТИВНА ПЛАТФОРМА ЗА ЕЛЕКТРОННО ОБУЧЕНИЕ</b> <i>Дуйгу Караолан Алтоп, Баръш Алтоп, Инанч Арън, Махмет Джем Айтекин, Муджахид Бенлиоулу, Юджел Сайгън</i>	89	<b>TOWARDS A MODULAR AND ADAPTIVE E-LEARNING PLATFORM</b> <i>Duygu Karaoğlan Altop, Baris Altop, İnanc Arın, Mehmet Cem Aytekin, Mucahid Benlioglu, Yücel Saygın</i>	2
3	<b>МЕТОДИ ЗА ИЗСЛЕДВАНЕ НА СЛУЧАЙНОСТТА НА НЕДВОИЧНИ ПСЕВДОСЛУЧАЙНИ ПОСЛЕДОВАТЕЛНОСТИ</b> <i>Антония Ташева, Боян Петров, Камен Стойков</i>	96	<b>METHODS FOR RANDOMNESS EVALUATION OF NONBINARY PSEUDORANDOM SEQUENCES</b> <i>Antoniya Tasheva, Boyan Petrov, Kamen Stoykov</i>	3



4	<b>СИСТЕМА ЗА ПУБЛИКУВАНЕ НА УЕБСАЙТОВЕ В СИСТЕМА ЗА СЪЗДАВАНЕ НА УЕБСАЙТОВЕ</b> <i>Димитър Димитров</i>	104	<b>A SYSTEM FOR WEBSITE PUBLISHING IN A WEBSITE BUILDING SYSTEM</b> <i>Dimitar Dimitrov</i>	4
5	<b>СИСТЕМА ЗА ОПТИМИЗАЦИЯ НА СНИМКИ В СИСТЕМИТЕ ЗА СЪЗДАВАНЕ НА УЕБСАЙТОВЕ</b> <i>Димитър Димитров</i>	112	<b>A SYSTEM FOR OPTIMIZATION OF IMAGES IN WEBSITE BUILDING SYSTEMS</b> <i>Dimitar Dimitrov</i>	5
6	<b>КЛАСИФИЦИРАНЕ НА МНЕНИЯ ОТ ПОТРЕБИТЕЛСКИ КОМЕНТАРИ НА АНГЛИЙСКИ ЕЗИК ЧРЕЗ АЛГОРИТМИ НА МАШИННОТО ОБУЧЕНИЕ</b> <i>Нели Арабаджиева-Калчева, Недялко Николов</i>	120	<b>CLASSIFICATION OF USER COMMENTS IN ENGLISH THROUGH MACHINE LEARNING</b> <i>Nely Arabadzhieva-Kalcheva, Nedyalko Nikolov</i>	6
7	<b>ИЗСЛЕДВАНЕ НА ПРИЛОЖИМОСТТА НА СЪЩЕСТВУВАЩИ МЕРКИ ЗА СХОДСТВО ЗА АНАЛИЗ НА ПОДОБИЕ МЕЖДУ ИЗОБРАЖЕНИЯ</b> <i>Мирослав Маринов, Ирена Вълва, Йордан Калмуков</i>	128	<b>COMPARATIVE ANALYSIS OF EXISTING SIMILARITY MEASURES EMPLOYED FOR CONTENT-BASED IMAGE RETRIEVAL</b> <i>Miroslav Marinov, Irena Valova, Yordan Kalmukov</i>	7
8	<b>НОВИ НЕВРОННИ МОДЕЛИ НА ЛРВНМ БАЗИРАНИ НА RELU АКТИВИРАЩИ ФУНКЦИИ</b> <i>Николай Дуков, Тодор Ганчев</i>	135	<b>NOVEL RELU-BASED NEURON MODELS FOR THE LRPNN</b> <i>Nikolay Dukov, Todor Ganchev</i>	8
9	<b>ОБРАБОТКА НА ТРИМЕРНИ ОБЕКТИ СЕГМЕНТИРАНИ ОТ МЕДИЦИНСКИ ИЗОБРАЖЕНИЯ</b> <i>Гергана Спасова, Николай Дуков, Кристина Близнакова</i>	143	<b>PROCESSING OF THREE- DIMENSIONAL OBJECTS SEGMENTED FROM PATIENT MEDICAL IMAGES</b> <i>Gergana Spasova, Nikolay Dukov, Kristina Bliznakova</i>	9
10	<b>ИЗСЛЕДВАНЕ НА АЛГОРИТМИ ЗА ОБХОЖДАНЕ НА ОБЕКТИ, РАЗПОЛОЖЕНИ В ЗОНА ЗА СИГУРНОСТ</b> <i>Илиян Бойчев</i>	151	<b>RESEARCH OF ALGORITMS FOR OBSERVATION THE OBJECTS IN A SECURE ZONE</b> <i>Iliyan Boychev</i>	10
11	<b>МОДЕЛИРАНЕ НА ХАРАКТЕРИСТИЧНА МЕЛОДИЯ НА ПТИЧИ ВОКАЛИЗАЦИИ</b> <i>Пролет Денева, Тодор Ганчев</i>	156	<b>MODELING OF CHARACTERISTIC MELODY OF BIRDS VOCALIZATIONS</b> <i>Prolet Deneva, Todor Ganchev</i>	11
12	<b>МЕТОД ЗА МОДЕЛИРАНЕ НА ТУМОРИ НА МЛЕЧНА ЖЛЕЗА</b> <i>Галя Господинова</i>	161	<b>AN APPROACH OF COMPUTATIONAL GENERATION OF IRREGULAR BREAST MASSES</b> <i>Galya Gospodinova</i>	12

# СЕКЦИЯ 1

## КОМПЮТЪРНИ СИСТЕМИ И МРЕЖИ

*Пета научна конференция с международно участие  
"Компютърни науки и технологии"  
28 – 29 септември, 2018 г.  
Варна, България*



*Fifth International Scientific Conference  
Computer Sciences and Engineering  
28 – 29 September, 2018  
Varna, Bulgaria*

## SECTION 1

### COMPUTER SYSTEMS AND NETWORKS

# TWO-FINGER TOUCH INTERFACE FOR WEARABLE DEVICES PROTOTYPE MODEL

Yuri Dimitrov

**Abstract:** Due to the wearable devices' small form factors (display size and controls) device interfaces development is challenging. Research in this paper proves that it is possible basic interface actions to be recognized using two-finger touch on the touch sensitive device bezel (the surface which is around the display) approach. The device interface algorithm based on the study results has been designed. An experimental model including hardware and software which performs these interface actions has been also build in the scope of the described in to the paper research.

**Keywords:** Wearable, Smartwatch, Touch, Interface

## 1. Introduction

Despite the fact that wearable devices are becoming very popular during the last few years [1], the user-device interaction techniques remain the same as those in the classic electronic watches - buttons, and in the smartphones - touch screen gestures. Touch based interfaces, being efficient and intuitive used in smartphones, when performed over a small surface like a smartwatch display, often face the problem that, due to the small sized interface controls, for the user is difficult to aim and push them correctly. This is defined by the researchers like "fat finger" problem [2]. Additionally, due to the line of sight obstacle caused by their own finger, they can't see what they push and the visual confirmation of the taken action.

Trying to solve these issues some of the leading vendors in the smartwatch area, and to make their devices more usable, have implemented in their recent models interfaces like rotating digital crown [3], touch sensitive rim in the outer part of the circle display [4], rotating and pushing bezel [5]. All of them are implemented in attempts the above mentioned "fat finger" problem to be solved. We consider that the alternative interfaces of wearable devices and specifically those which allow interfacing without user sight interruption have to be further researched.

## 2. Related work

There are related researches as [6] where the authors evaluates the interaction at the outer side of round wearable devices. The major difference between this research and the current one is that here have been evaluated the finger movement range while the other authors evaluate how precise is the aiming of the touches. There are also differences in the experimental model - the used in this research is 44 mm in diameter and is much more close the standard smartwatches while the other model is relatively bigger - 60 mm in diameter. Also the sensors' position place is not the same - in the model used for this research they are on the bezel, which is a much more accessible part of wearable devices, while on the mentioned model they are on the edge.

## 3. Study

The goal of the Study is to collect data what are the user finger movement ranges, when they execute "down" and "up" menu selection actions (scrolls) over the touch sensitive device bezel. For successful results from the study will be counted results where at least 2 new touched or untouched consecutive sensors in the finger movement direction while "down" and "up" actions are performed.

**Equipment:**

For the Study a 3D model of a Smartwatch, which was designed and printed for previous research, was used. On the bezel of the model 12 capacitive sensors mounted. The distance between each two sensors was also 6 mm. The experimental model was driven by an Arduino Mega 2560 computer with a 12 capacitive ports MPR121 extension board connected to it. The 12 capacitive sensors were connected to the MPR121 extension board sensors' ports. To each sensor was assigned an ID from 1 to 12 corresponding to the 12-hour standard watch hour mark place of each sensor. Specially developed for the research purposes software in the Arduino language was used for the experiment. The software detected the starts (touching contact) and the ends (releasing) of the touches made by the participant's fingers independently for each sensor. The Arduino Mega 2560 computer was connected to a notebook via USB interface and sent the data to the Arduino IDE Serial Monitor. The raw data was then transferred to MS Excel for further processing.

**Group:**

The test group consisted of 10 people (male 8, female 2), aged 26-48 (average 33.9), all right handed, all volunteers.

**The process:**

The instructions, given to each participant, were to wear the model on his left hand. After that to raise his/her left hand in position allowing him/her to see the device display. To touch the bezel of the device with two fingers - thumb and one of the other fingers on his/her leading hand. After that to scroll three times down and three times up over the device bezel with his/her pointing (or other finger) while remain the touch by his/her thumb. After each attempt the participant was instructed to go back to a relaxed sitting position. Each participant made three consequential attempts for both "down" and "up" scrolls. On Fig.1 is shown the study process and the experimental model. The data from all three attempts / six sets of movements was recorded in a file.



**Fig. 1.** Experimental model and study process

**Study data:**

A sample of the collected raw experiment data is shown in Table 1, where is represented a data set for one User for one performed action - User 2, action "down". The "1" in rows "Touched" shows if the particular sensor has been touched during the user finger scroll over device bezel, and in rows "Released" shows if the already touched sensor during to action perform has been released.

**Table 1.** Sample of Study raw data per User 2 per one Operation - “down”

User	Sensors	1	2	3	4	5	6	7	8	9	10	11	12
1	Touched			1	1								
	Released	1	1										
1	Touched			1	1	1							
	Released		1	1									
1	Touched			1	1	1							
	Released		1	1	1								

The summary data from all experiments were for each recorded “down” and “up” where the number of touched next sensor in the scroll/gesture movement direction as well as the number of the released sensors during the same finger movement.

**Table 2.** Summary data of number of touched and released sensors during the “down” and “up” operations perform during the study experiment

User/ Attempt	Sensors	Number DOWN	Number UP	User/ Attempt	Sensors	Number DOWN	Number UP
1/1	Touched	3	2	6/1	Touched	2	2
	Released	3	2		Released	1	1
1/2	Touched	3	3	6/2	Touched	3	3
	Released	3	3		Released	3	2
1/3	Touched	3	3	6/3	Touched	2	1
	Released	3	2		Released	2	2
2/1	Touched	2	2	7/1	Touched	2	2
	Released	2	2		Released	2	1
2/2	Touched	3	3	7/2	Touched	2	2
	Released	2	2		Released	3	2
2/3	Touched	3	3	7/3	Touched	3	3
	Released	3	2		Released	3	3
3/1	Touched	2	2	8/1	Touched	3	2
	Released	2	2		Released	2	2
3/2	Touched	3	2	8/2	Touched	2	2
	Released	2	1		Released	2	2
3/3	Touched	2	2	8/3	Touched	3	2
	Released	1	1		Released	3	2
4/1	Touched	3	3	9/1	Touched	2	2
	Released	3	2		Released	2	2
4/2	Touched	2	2	9/2	Touched	2	2
	Released	1	1		Released	2	2
4/3	Touched	2	2	9/3	Touched	3	2
	Released	2	2		Released	2	2
5/1	Touched	3	3	10/1	Touched	2	2
	Released	4	2		Released	1	1
5/2	Touched	2	2	10/2	Touched	3	3
	Released	3	3		Released	2	2
5/3	Touched	2	2	10/3	Touched	3	1
	Released	2	2		Released	2	2

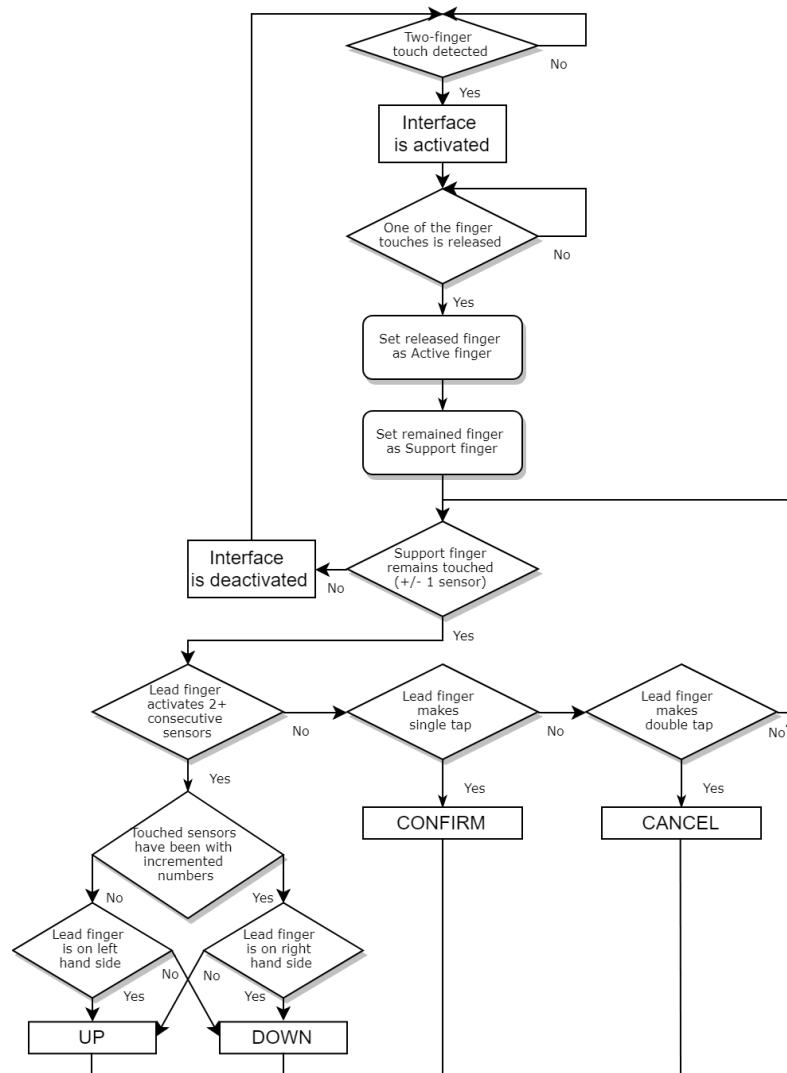
### Study results:

The Average and Mode values of the study summary data for both “down” and “up” user cases are presented in Table 3.

**Table 3.** Average and Mode values from the Study data

Direction	Sensors	Average	Mode	Minimal number (T or R)
DOWN	Touched	2.50	3	2
	Released	2.67	2	
UP	Touched	2.23	2	2
	Released	1.90	2	

These results confirm that it is possible “down” and “up” scrolls over the touch sensitive device bezel with 12 touch sensors to be recognized as the Minimal number of the touched or released sensors in each of all 60 attempts have been made is 2. It is expected that the model, which is going to be built following this approach, to work stable as the Mode value is at least 2 and the Average values are higher than 2 (excluding 1.90 for released in “up” scroll).



**Fig. 2.** Block diagram of the software part of the model

## 4. Prototype model

The goal is a prototype model of wearable device interface to be developed, based on the results from the paper study. The inputs of the model are user's two-finger touches over the device touch sensitive bezel, which have to activate the interface, to perform some set of the four different actions and to deactivate the interface. The output of the model are four commands based on user touches: "down" and "up" scrolls for menu item selection, "confirm" tap (confirmation/enter action) and "cancel" double tap (escape to the previous menu action).

The prototype model has two parts - hardware one and software one. The used hardware equipment for the prototype model development is exactly the same as the used one during the paper study experimental process. The software part has to operate following the process shown on diagram on Fig. 2.

### Results:

In order the prototype model to be tested a sets of actions have been executed and the results have been compared with the input set. After the tests it has been proven that the prototype model works as it supposed. A sample from test process output when the input had been set to "down"- "down"- "up" is presented on Fig. 3

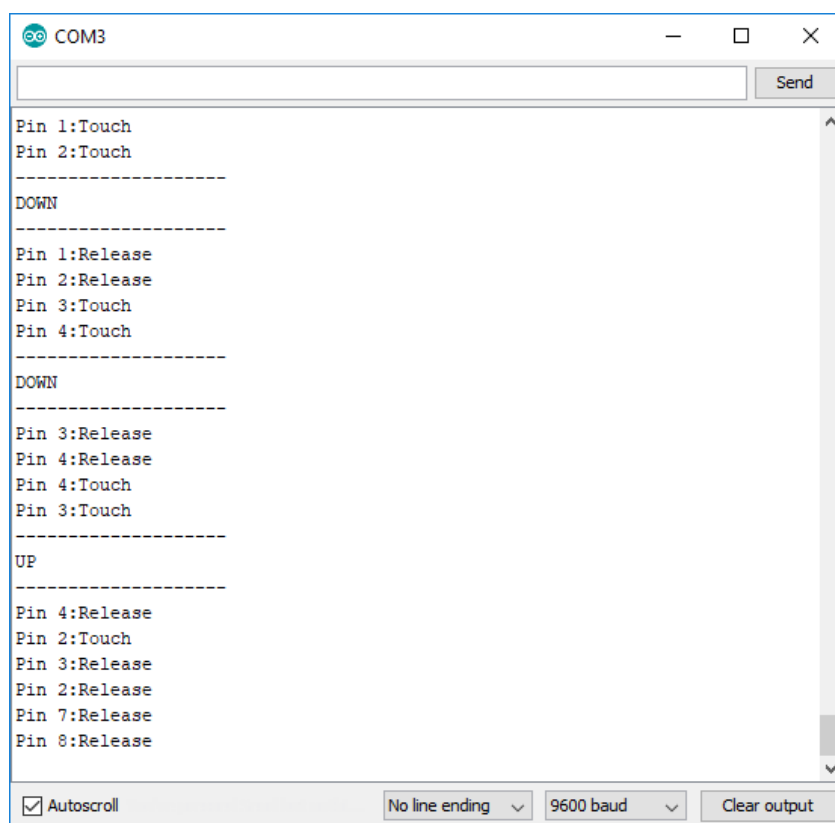


Fig. 3. Model output

## 5. Conclusion

The paper results from the experimental study clearly show that interfaces, based on movement of one finger over the touch sensitive device bezel, and single and double taps, while the other finger is used for interface activation, deactivation and gesture support, are possible and applicable for wearable device interface development.



The achieved functionalities of the prototype model prove that the experimentally obtained results could be used for the development of a fully functional prototype model, both using one and the same hardware model.

## 6. Further work

Based on the paper results and conclusions, in order the efficiency and robustness of the two-finger touch interface over the wearable device touch sensitive bezel to be measured and evaluated, a comparative analyse between this interface and tradition touch screen and push buttons interfaces should be performed.

The software part of the prototype model could be improved by the algorithm further development in order more actions/commands to be executed (more gestures to be recognized by the software) and by adding of the management of the time delays and cycles periods of reading data from the capacitive sensors.

The hardware part of the prototype model could be also improved by increasing the number of used sensors in order a higher movement resolution to be achieved as well as experimenting with different sensors' sizes or touch technologies for better gesture recognition.

## References

- [1] IDC, "International Data Corporation (IDC) Worldwide Quarterly Wearable Device Tracker - Q3 2017, <https://www.idc.com/getdoc.jsp?containerId=prUS43260217>" - last visited on August 24, 2018.
- [2] K. A. Siek, Y. Rogers, and K. H. Connelly, "Fat Finger Worries:\nHow Older and Younger Users\nPhysically Interact with PDAs," pp. 267–280, 2005.
- [3] Apple, "<https://www.apple.com/watch/>" - last visited on August 24, 2018.
- [4] Misfit, "<https://misfit.com/misfit-vapor>" last visited on August 24, 2018.
- [5] Samsung Electronics, "<http://www.samsung.com/uk/wearables/gear-s3/>" last visited on August 24, 2018.
- [6] R. Darbar *et al.*, "Interaction on the Edge: Offset Sensing for Small Devices," *Proc. 2016 CHI Conf. Hum. Factors Comput. Syst. - CHI '16*, vol. 13, no. 4, pp. 169–178, 2016.

### For contacts:

Eng. Yuri Dimitrov  
Computer Science and Engineering Department  
Technical University of Varna  
E-mail: [yndimitrov@gmail.com](mailto:yndimitrov@gmail.com)



# СИСТЕМА ЗА СИМУЛИРАНЕ НА МЕХАНИЗМА НА ИЗГРАЖДАНЕ НА НАЙ-ПРЕДПОЧИТАНИЯ ПЪТ ПРИ 6LoWPAN ТЕХНОЛОГИЯ

Айдън М. Хъкъ, Русен В. Василев

**Резюме:** В този доклад се предлага симулационен модел за анализиране ефективността на механизма за изграждане на най-добър маршрут при 6LoWPAN мрежи. Представен е механизъм за изграждане на най-добър маршрут, който проследява състоянието на всеки изграден маршрут, за да гарантира предаването на спешна информация и увеличава броя на опитите за повторно предаване на пакети.

**Ключови думи:** 5G, IoT, 6LoWPAN, симулация

## System for simulating the construction of the best route in 6LoWPAN technology

Aydan M. Haka, Rusen V. Vasilev

**Abstract:** This paper proposes a simulation model which analyses the best route construction mechanism for 6LoWPAN networks. The mechanism is for building the best route and tracking the status of every available option to ensure the transmission of emergency information as well as recording the number of attempts to retransmit packets.

**Keywords:** 5G, IoT, 6LoWPAN, simulation

### 1. Увод

В доклада си Ericsson [1] съобщава, че до 2023 г. се предвиждат над 30 милиарда свързани устройства, от които 20 милиарда ще бъдат свързани с Internet of Things (IoT). Мрежите от IoT устройства включват свързани автомобили, машини, сензори, потребителска електроника, терминали, мобилни устройства и др. В зависимост от това какви устройства са включени в IoT мрежата и каква е тяхната функционалност, комуникацията между тях се осъществява на базата на съответния протокол. Известни са редица IoT протоколи [2] за комуникация като Z-Wave, Thread, ZigBee, Bluetooth и др., които определят използваните механизми за комуникация, поддържани скорости, модулационни схеми и т.н. Свързаността на тези устройства не се ограничава в локалната мрежа, тя е част от глобалната мрежа, като клетъчната свързаност се очаква да се осигури основно от LTE (олекотената версия LTE-M1, която е част от 3GPP release 12 и 13) [3] и 5G (SDR-based Base Stations) [4].

Концепцията IoT представлява безжична мрежа от свързани устройства - главно безжични сензори с различни възможности. Една от технологиите, изграждащи IoT, е IPv6 over Low-Power Wireless Personal Area Networks (6LoWPAN), базирана на комуникационния стандарт IEEE 802.15.4 [5]. Тя предоставя адаптивен слой, който позволява предаването на IPv6 пакети чрез малки Link Layer фреймове [6]. 6LoWPAN сензорните устройства могат да бъдат използвани за проследяване на температура, влажност, аларми при пожар, измерване на водно ниво, измерване на вибрации, домашна сигурност, контрол и автоматизирани системи [7]. В зависимост от средата и ситуацията, в която се използват тези сензори, може да бъде наложително проследяването на важна или жизнено необходима информация. Нуждата от следене на такива критични параметри налага използване на механизми за ускорено предаване на информационните данни по мрежата и осигуряване предимство на спешната информация. Това може да се постигне чрез приоритизиране на спешния трафик, изграждане на най-добър маршрут за предаване на пакетите и следене състоянието на тези маршрути.

Тази статия представя разработена среда, чрез която може да се симулира изграждането на 6LoWPAN мрежа, в която се реализира приоритизиране на предавания трафик според изискванията на мрежата, изграждане и избор на най-добър маршрут от точка до точка и постоянно проследяване състоянието на изградените маршрути.

## 2. Същност на използвания Emergent Direct Mode

В [8] са описани две основни схеми за маршрутизиране при 6LoWPAN – Mesh Under и Route Over, като са описани предимствата и недостатъците им. Тук се предлага и Emergent Direct Mode режим за свързване, който преодолява недостатъците на разгледаните маршрутизиращи схеми.

На базата на това изследване предложеният симулатор имплементира в себе си Emergent Direct Mode режим за свързване, състоящ се от Hassle Free Route механизъм за маршрутизиране и Adaptive Retry механизъм за подобряване на повторното изпращане. Този механизъм за маршрутизиране позволява приоритизирането на пакети, пренасящи важна или жизнено необходима информация според изградената 6LoWPAN мрежа. Според Emergent Direct Mode схемата всеки пакет получава нова dispatch хедър стойност за 6LoWPAN хедъра на пакета, която идентифицира схемата за маршрутизиране. Когато пристигне спешен пакет, той се фрагментира с dispatch хедъра, за да бъде третиран като специален.

След като е идентифициран Emergent Direct Mode режимът, предпочитаният маршрут се изгражда чрез Hassle Free Route механизъм и механизъм за контрол на повторното предаване се обновява с Adaptive Retry Mechanism.

Според Hassle Free Route механизма в 6LoWPAN мултихоп мрежите най-добрият маршрут се избира според това дали е най-краткият и се добавя автентична стойност за всеки маршрут, което определя приоритета му. Автентичната стойност се запазва в маршрутната таблица на възлите и тя може да бъде:

- отрицателно число – информира за това, че по маршрута са реализирани твърде много неуспешни предавания;
- стойност „0“ – показва, че маршрутът все още не е оценен;
- положителна число – маркира, че маршрутът е безпроблемен. По-голямата стойност означава, че маршрутът е по-успешен от другите, като по този начин се обозначава броя на успешно предадените фрагменти по определен маршрут.

При всяко успешно предаване автентичната стойност се увеличава с „1“, а в случай на неуспешно предаване се намалява с „1“. Когато пристигне фрагмент с Emergent dispatch хедър, той се препраща по маршрута с най-висока автентична стойност. Тези фрагменти са с по-висок приоритет и се предават по най-предпочитания маршрут.

Emergent Direct Mode схемата използва и Adaptive Retry Mechanism за осигуряване на повторно предаване, гарантиращо, че спешните фрагменти няма да отпаднат по маршрута. При този механизъм стойността за повторно предаване се разделя на две части: основна стойност и променлива стойност в съотношение 2:1. Ако контролната стойност за механизма на повторно предаване е 15, то основната стойност ще е 10, а променливата стойност 5. Основната стойност представлява стойността за контрол на препредаването. Винаги, когато се открие Emergent Direct Mode, се отчитат както основната стойност, така и променливата стойност. По този начин могат да се реализират допълнителни препредавания, което осигурява, че в рамките на претоварена мрежа фрагментите с Emergent dispatch хедър ще бъдат предадени със сигурност.

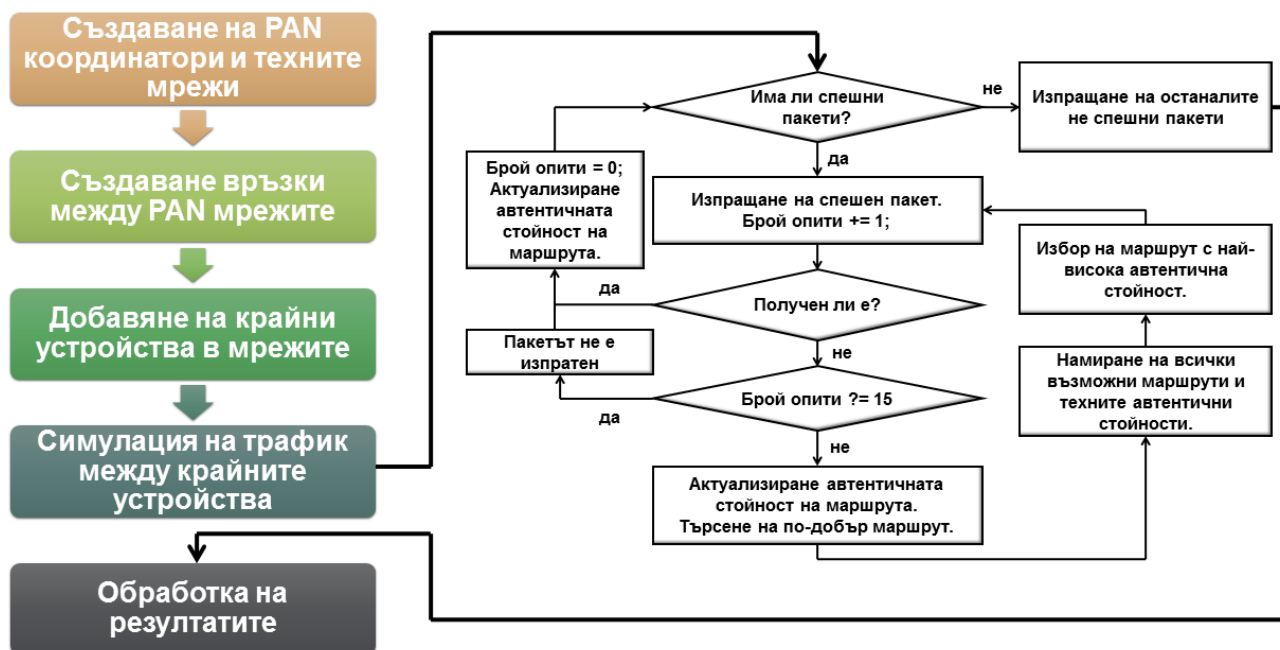
За експериментални цели основната стойност за повторно предаване Rconst се приема за 10. Adaptive Retry Mechanism (Rnew) се предлага като:

$$R_{new} = R_{const} + SF \bmod 5 \quad (1)$$

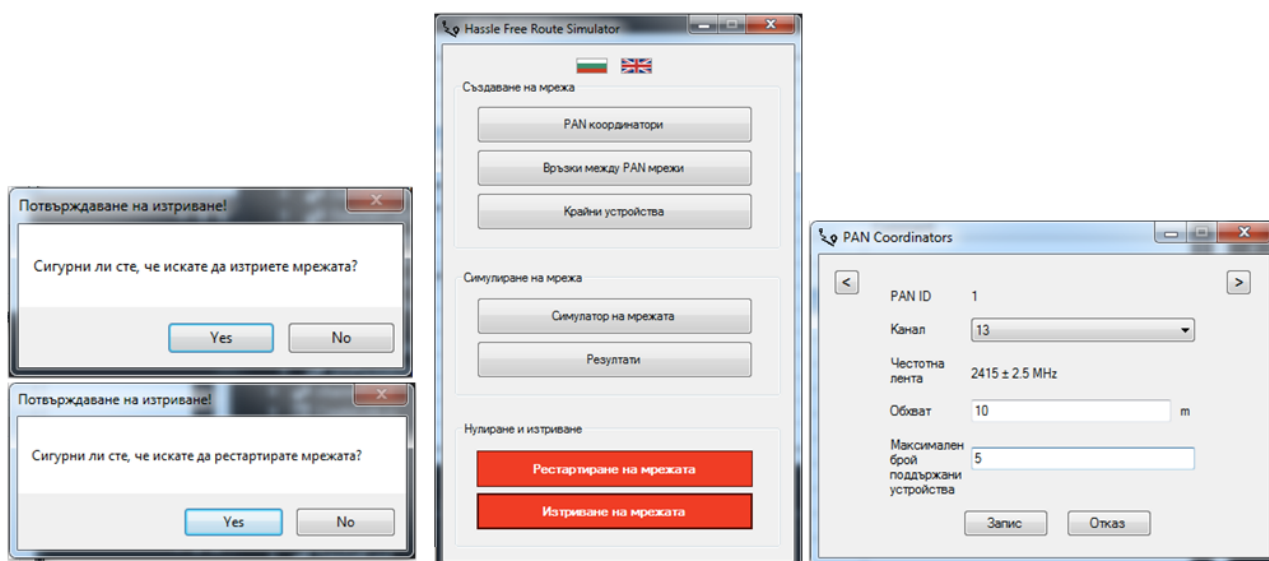
където SF (Significant Factor) се определя от възникващия приоритет на данните пакети. Този механизъм осигурява стойността за повторно предаване да е между 5 и 15 опита, където 15 е максималният брой опити.

### 3. Симулационна среда и експериментални изследвания

За анализиране на прегледания в настоящото изследване алгоритъм за приоритизиране на трафика, изграждане и проследяване състоянието на възможните маршрути при 6LoWPAN мрежи, се създава модел на 6LoWPAN мрежа с няколко свързани координатора. Процесът на създаване на 6LoWPAN мрежа и симулирането на работата на използваната Emergent Direct Mode схема е представено постъпково на фигура 1.

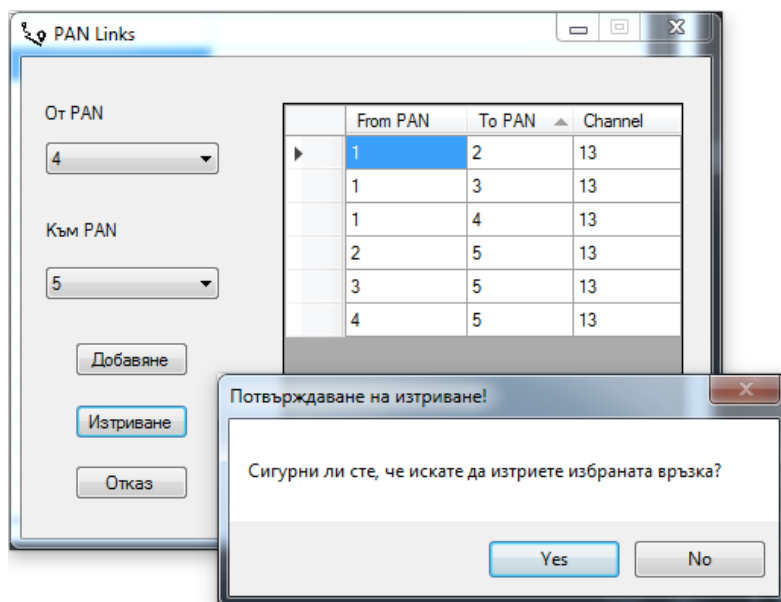


Фиг. 1. Стъпки при изграждане на 6LoWPAN мрежа



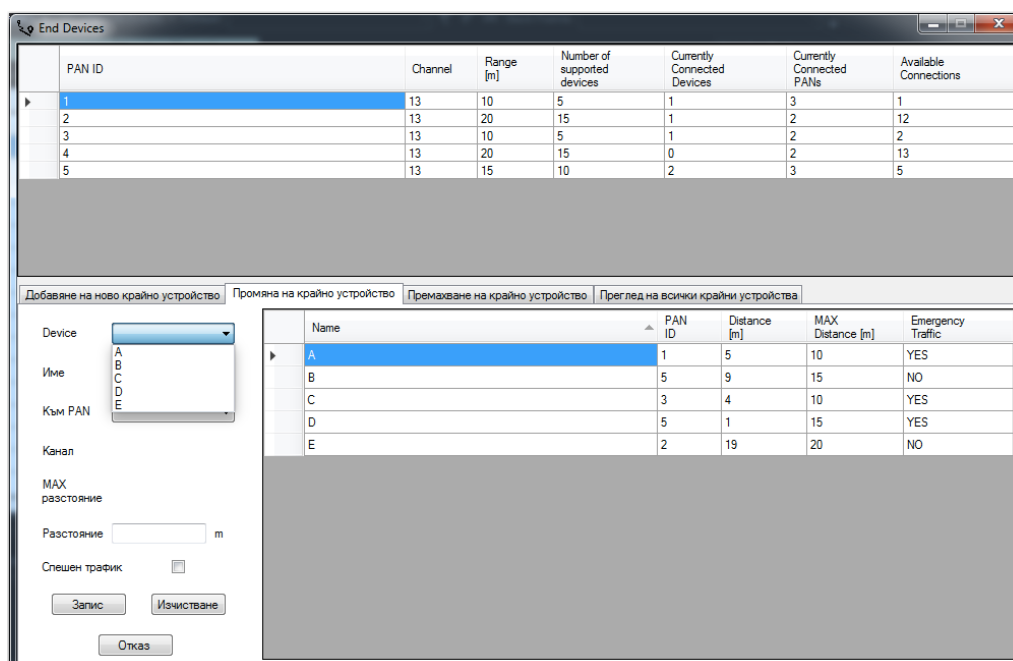
Фиг. 2. Създаване на 6LoWPAN мрежа

Използваното програмно средство за създаване на симулационния продукт е Microsoft Visual Studio 2012 и езикът за програмиране C#, защото позволява използването на събитийно-ориентирана среда за разработка. Изградена е база от данни за съхранение на данните от отделните експерименти за всеки 6LoWPAN координатор и свързаните към тях крайни устройства. Реализираният симулатор имплементира в себе си мултиезичен интерфейс, като до момента са включени английски и български език. Създаването на 6LoWPAN мрежата се осъществява чрез интерфейса, представен на фигура 2. Между PAN координаторите и техните мрежи се създават връзки, по които се предават пакетите, докато достигнат желаната дестинация (фигура 3). Изградените връзки са двустранни и се създават еднократно.



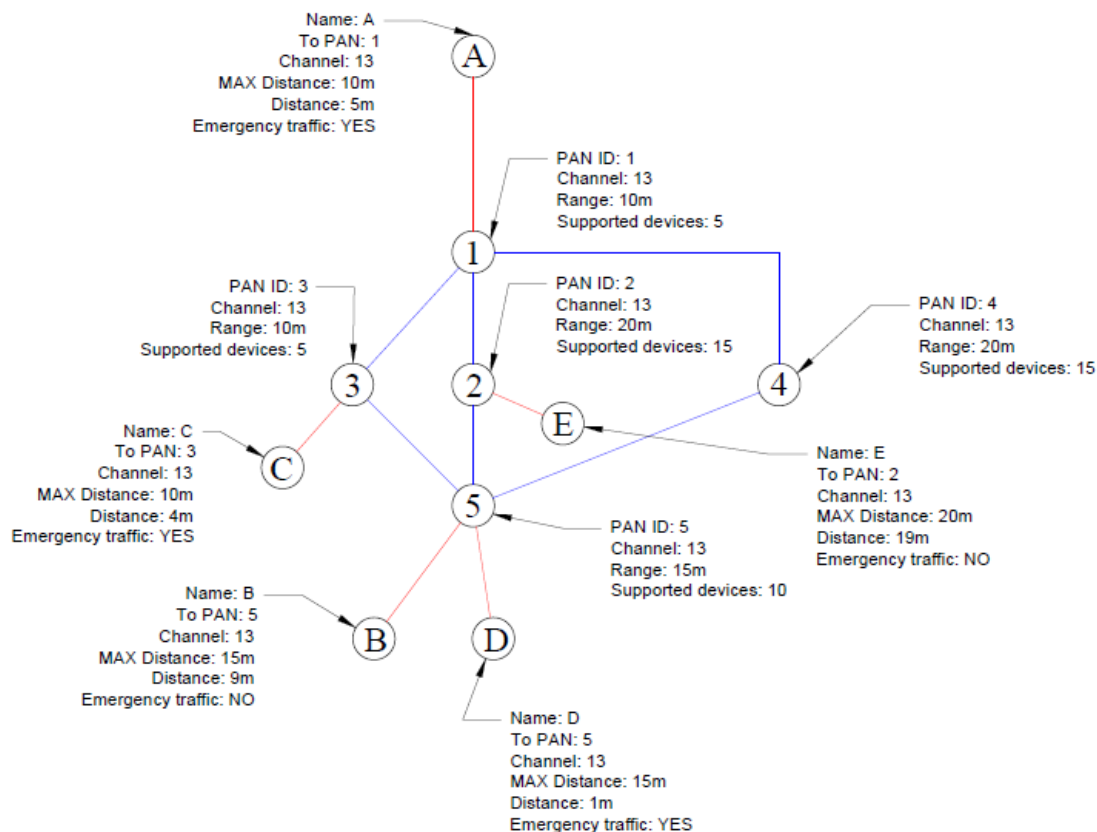
**Фиг. 3.** Създаване на връзки между 6LoWPAN координаторите и техните мрежи

За всяко крайно устройство или координатор има възможност за актуализация на данните, както е показано на фигура 4.



**Фиг. 4.** Актуализиране на данните за координаторите и крайните устройства

След като е въведена информацията за координаторите, връзките между тях и крайните устройства към всеки координатор, според представената топология на фигура 5, се реализира симулирането на разгледания Emergent Direct Mode. Симулаторът приоритизира спешния трафик, ако има такъв, и го предава по най-добрия маршрут, избран според Hassle Free механизма, като за всеки маршрут се проследява автентичната му стойност и се преценява колко е надежден (фигура 6). Преди реализиране на предаването всички пакети се добавят към опашката за изпращане, където се приоритизират и предават по избрания маршрут.



Фиг. 5. Топология на 6LoWPAN мрежа

The screenshot shows the Network Simulator interface with the following components:

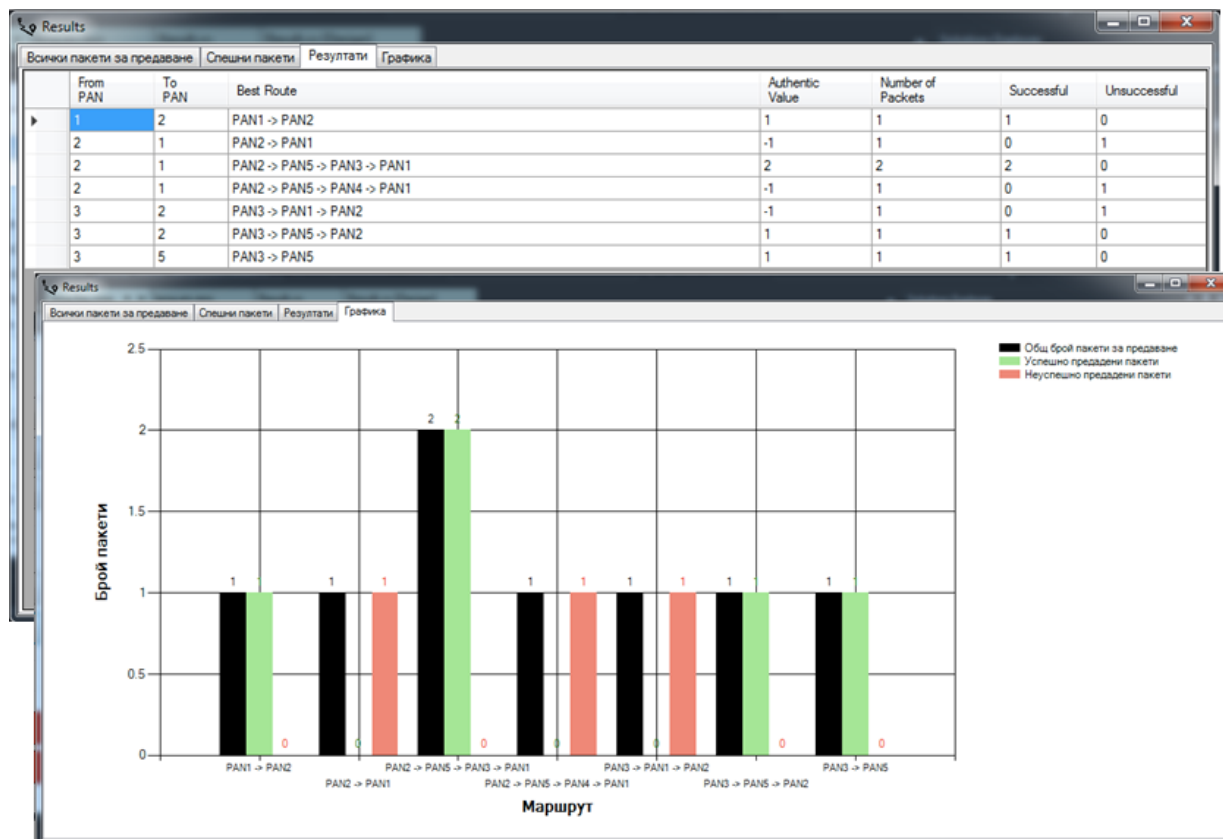
- Источник и дестинация:**
  - От устройство: C
  - До устройство: E
  - Спешен трафик: **ДА**
  - Buttons: Приоритизиране!, Добавяне на пакета към опашката, Отказ на текущия избор
- Възможни маршрути:**

From Device	To Device	Available Routes	Authentic Value
C	E	PAN3 -> PAN1 -> PAN2	0
C	E	PAN3 -> PAN1 -> PAN4 -> PAN5 -> PAN2	0
C	E	PAN3 -> PAN5 -> PAN4 -> PAN1 -> PAN2	0
C	E	PAN3 -> PAN5 -> PAN2	0
- Най-предпочитан път:** PAN3 -> PAN1 -> PAN2
- Опашка с пакети:**

Emergency Packet	From Device	To Device	From PAN	To PAN	Best Route	Authentic Value	Status
Yes	C	E	3	2	PAN3 -> PAN1 -> PAN2	0	
No	B	A	5	1	PAN5 -> PAN4 -> PAN1	0	
Yes	C	D	3	5	PAN3 -> PAN5	0	
Yes	E	A	2	1	PAN2 -> PAN5 -> PAN3 -> PAN1	1	

Фиг. 6. Избор на най-добър маршрут

На фигура 7 са показани резултатите от приложението на разгледаните по-горе механизми за избор на най-предпочитан маршрут при 6LoWPAN технология. Графиката представя общия брой пакети за предаване по всеки маршрут, съответно колко са били предадени успешно и колко - не, според което се оценява автентичната стойност на съответния маршрут.



Фиг. 7. Табличен и графичен вид на избраните маршрути и предадените пакети

#### 4. Заключение

В доклада е представен механизъм за приоритизиране на спешен трафик и избор на най-предпочитан маршрут в 6LoWPAN мрежа. За изследване на разгледания механизъм е разработена симулационна среда. Настоящият модел може да се ползва в учебния процес в дисциплини, свързани с изследване на метода за избор на най-добър маршрут в безжични сензорни мрежи, базирани на 6LoWPAN технологията.

Цел на бъдещата работа е да се разшири функционалността на създадения симулатор, така че да позволява визуализиране на матрицата на предаване и реализиране на мобилност на крайните устройства между 6LoWPAN координаторите с цел изследване на QoS параметрите при тези мрежи.

#### Литература

- [1]. Ericsson Mobility Report. November 2017. <https://www.ericsson.com/en/mobility-report/reports/november-2017> . Last visit on 02.07.2018
- [2]. The Complete List Of Wireless IoT Network Protocols. 08 February 2016. <https://www.link-labs.com/blog/complete-list-iot-network-protocols> . Last visit on 08.07.2018
- [3]. Three 3GPP IoT Technologies To Get Familiar With. 29 November 2017. <https://www.link-labs.com/blog/3gpp-iot-technologies> . Last visit on 08.07.2018



- [4]. IoT and 5G: spectrum and networking. 2015. <http://iot-week.eu/wp-content/uploads/2015/07/2-IoT-and-5G-spectrum.pdf> . Last visit on 08.07.2018
- [5]. IEEE Standard for Local and metropolitan area networks – Part 15.4: Low-Rate Wireless Personal Area Networks(LR-WPANs). 5 September 2011. [http://ecee.colorado.edu/~liue/teaching/comm\\_standards/2015S\\_zigbee/802.15.4-2011.pdf](http://ecee.colorado.edu/~liue/teaching/comm_standards/2015S_zigbee/802.15.4-2011.pdf). Last visit on 03.07.2018
- [6]. IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals. August 2007. [https://datatracker.ietf.org/doc/rfc4919/?include\\_text=1](https://datatracker.ietf.org/doc/rfc4919/?include_text=1) . Last visit on 03.07.2018
- [7]. Design and Application Spaces for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs). April 2012. <https://tools.ietf.org/html/rfc6568>. Last visit on 03.07.2018
- [8]. Saha, R., Bhunia, S. S., Mikherjee, N. Enhancing 6LoWPAN scheme to support priority driven routing. //2015 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), 15-18 Dec. 2015

**За контакти:**

ас. инж. Айдын М. Хъкъ  
катедра „Компютърни науки и технологии”  
Технически университет-Варна  
E-mail: aydin.mehmed@tu-varna.bg

инж. Русен В. Василев  
катедра „Компютърни науки и технологии”  
Технически университет-Варна  
E-mail: rusen.ru@gmail.com



# СРАВНИТЕЛЕН АНАЛИЗ НА МЕХАНИЗМИ ЗА ПРИОРИТИЗАЦИЯ В LTE SCHEDULER

Айдън М. Хъкъ

**Резюме:** В този доклад е реализиран сравнителен анализ на три приоритизационни механизма за LTE Scheduler, разработен от автора, Myo и Akyildiz. Сравнението е реализирано на база на резултатите, получени спрямо трите алгоритъма за VoIP и nonGBR услуга. Сравняваните параметри са Throughput, Packet Delivery Ratio и Delay. Резултатите от извършения анализ водят до заключението, че по отношение на сравняваните параметри, предложението от автора приоритизационен механизъм има предимство при обслужване на мобилни абонати, движещи се с високи скорости.

**Ключови думи:** LTE, Scheduler, Приоритизация

## Comparative analysis of prioritization mechanisms in LTE Scheduler

Aydan M. Haka

**Abstract:** This paper represents a comparative analysis of three prioritization mechanisms for the Scheduler of LTE, developed by the author, Myo and Akyildiz. The comparison is based on the results obtained with the three algorithms for VoIP and nonGBR service. Compare parameters are Throughput, Packet Delivery Ratio, and Delay. The results of the analysis lead to the conclusion that, with respect to the parameters being compared, the prioritization mechanism of the author has the advantage of servicing mobile subscribers with high speeds.

**Keywords:** LTE, Scheduler, Prioritization

### 1. Увод

Системата за клетъчна комуникация, Long Term Evolution (LTE), се налага като преобладаваща технология на световния пазар [1]. Тя е част от Evolved Packet System (EPS), разработена от 3GPP. EPS включва няколко технологии, част от които са Orthogonal Frequency Division Multiplex (OFDM) схема за достъп в Downlink посока и Multiple Inputs Multiple Outputs (MIMO) технология на антените за комуникация [2].

Широкото разгръщане на LTE мрежите по света е предпоставка за увеличаване броя на клетъчните абонаменти [3, 4], като това ще увеличи обема на трафика, предаван между базовите станции (eNodeB) и абонатите (User Equipment - UE). С увеличаване обема на предавания трафик ще се увеличат и изискванията за качеството на.

Обслужването на UEs се реализира от Scheduler-а в eNodeB. Той трябва да разпредели радиоресурсите ефективно, като по този начин да гарантира поддържане на голямо разнообразие от услуги и увеличаване производителността на системата. Това може да се постигне от Scheduler-а чрез приоритизиране заявките на обслужваните от eNodeB абонати.

Тази статия представя сравнителен анализ между алгоритъма за приоритизация в [5] с предложените в [7] и [8] механизми с цел изследване качеството на услугите обслужване (Quality of Service - QoS) за свързаните абонати.

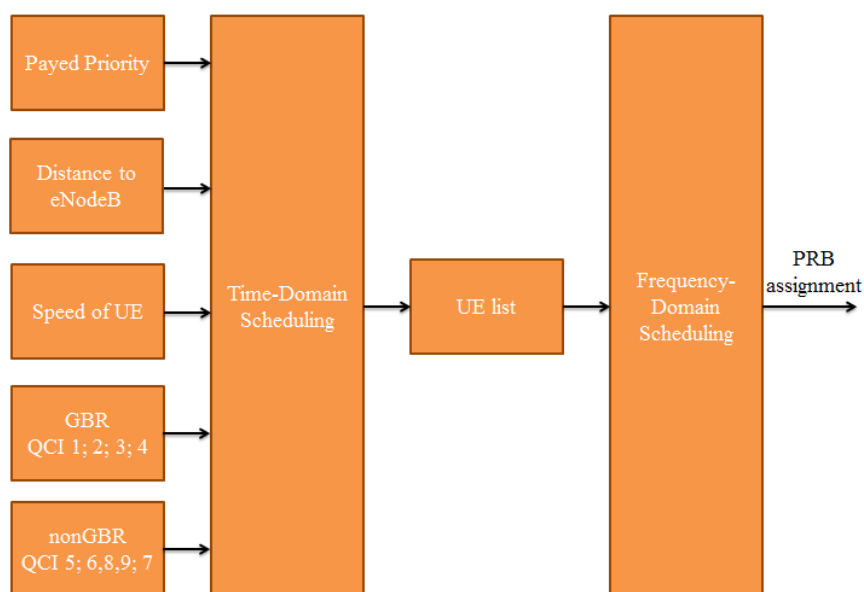
### 2. Планиране на LTE ресурсите

Планирането е метод за динамично разпределение на радиоресурсите по абонати, базирано на алгоритъма за планиране и приоритизация в LTE базовата станция. За разпределяне на ресурсите в Downlink посока при LTE се използва технологията OFDM/OFDMA [9], при която наличната честотна лента е разделена на паралелни

теснолентови подносещи честоти с разстояние от 15 KHz, независимо от общата широчина на честотната лента. Радиоресурсите се дефинират във времево-честотния домейн. Във времевия домейн времето е разделено по фреймове и всеки LTE фрейм се състои от десет последователни подфрейма или т. нар. TTI (Transmission Time Interval), всеки от които е с продължителност 1 ms, т. е. всеки LTE фрейм е с продължителност 10 ms. Всеки подфрейм се състои от два Timeslot-a, като всеки слот е с продължителност 0.5 ms. Отделните Timeslot се състоят от 7 OFDM символа при кратък цикличен префикс (6 символа за удължен цикличен префикс). В честотния домейн цялата налична честотна лента е разделена на подканални. Всеки подканал се състои от 12 последователни и еднакво разположени подносещи с разстояние между тях от 15 KHz, т. е. всеки подканал е с широчина 180 KHz ( $12 \cdot 15$  KHz). Времево-честотният ресурс обхващащ един подканал от 12 подносещи в честотния домейн и един Timeslot от 0.5 ms във времевия домейн е известен като ресурсен блок (Resource Block - RB). Най-малката единица от RB, която съответства на една OFDM подносеща по време на един символ интервал, е известна като ресурсен елемент (Resource Element - RE). При планирането ресурсите се разпределят по двойки, известни като Physical Resource Block (PRB). PRB може да се определи като минималния брой ресурси, които могат да бъдат предоставени на UE за предаване на данни във всеки TTI.

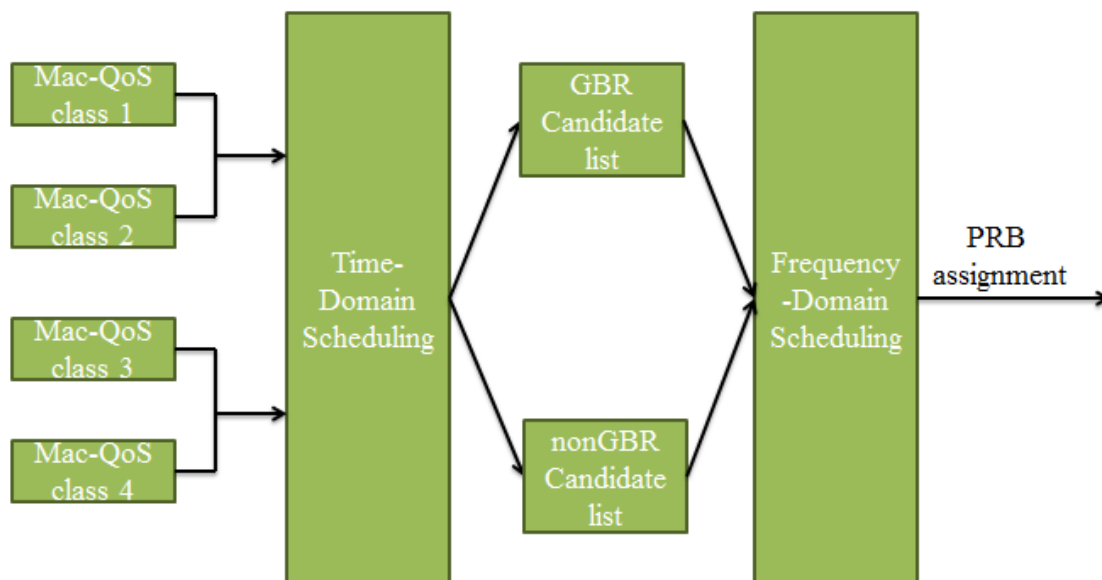
При LTE планирането и разпределението на ресурсите се реализира от eNodeB. За да се подобрят QoS за абонатите е необходимо да се разгледат различни приоритизационни механизми за Scheduler-a на eNodeB. Тук се разглеждат два алгоритъма за приоритизация на трафика по абонати:

- Предложеният в [5] алгоритъм спазва изцяло изискванията на стандарта за приоритизация, описани в [6]. Освен стандартните изисквания трафикът се приоритизира и според допълнителни параметри като заплащане за гарантирана услуга, което позволява на ISP да преприоритизира трафикът, а абонатите от своя страна да преценят за коя услуга имат нужда от гарантиран достъп. Този алгоритъм още предоставя предимство при обслужването на мобилни абонати, тъй като те са чувствителни спрямо качеството на обслужване. Разпределението на радиоресурсите по абонати зависи от приоритета - тези с по-висок приоритет получават повече ресурси от нископриоритетните UE. Тези заявки, за които няма достатъчно ресурс, се отлагат за обслужване в следващ фрейм.



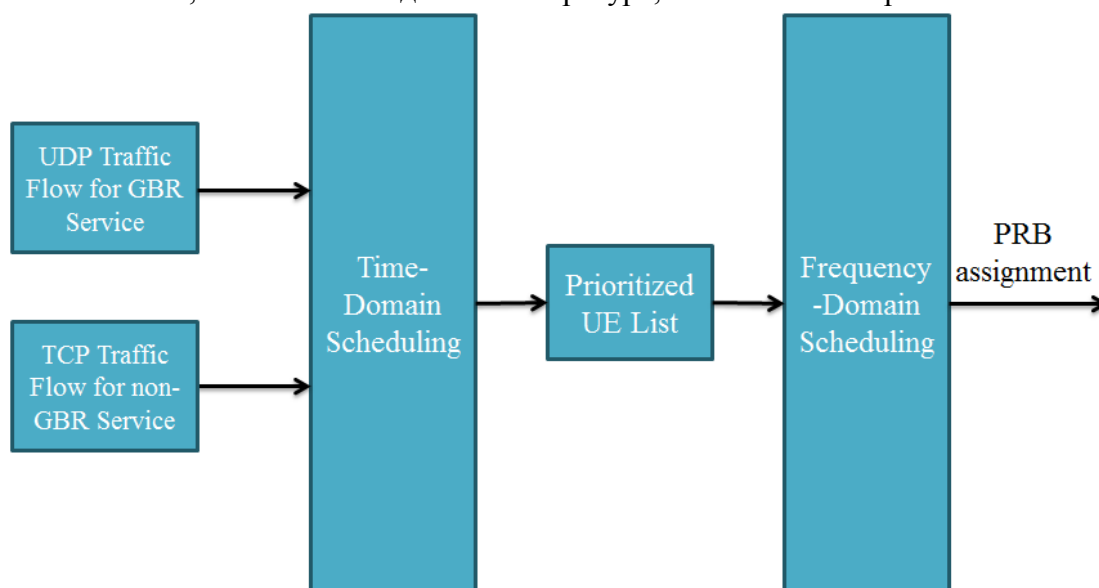
Фиг. 1. Блок-схема на предложената приоритизация

- Представеният в [7] QoS-Aware Proportional Fair (QAPF) Scheduler реализира приоритизацията, спазвайки частично стандарта, описан в [6]. Трафикът се приоритизира според типа на Bearer: GBR (Guaranteed Bit Rate) – с по-висок приоритет за услуги Voice over IP (VoIP) и Live Video Streaming; non-GBR (non-Guaranteed Bit Rate) – с по-нисък приоритет за услуги Video Streaming for buffer и Background traffic. За GBR приоритизацията основно се базира на Head of Line (HoL) забавянето, ако HoL забавянето надхвърля максималния бюджет за забавяне, пакетите се отхвърлят, докато при non-GBR зависи от приоритета на изисквания тип услуга по стандарт. Разпределението на ресурсните блокове става по равно за всеки абонат, като се започне с тези, изискали GBR услуга;



Фиг. 2. Блок-схема на QAPF приоритизация

- Предложеният в [8] алгоритъм не се базира на стандарта в [6]. Той позволява приоритетно обслужване на GBR услугите. Въпреки това, при по-голям процент изискани non-GBR услуги, те се обслужват с приоритет (при равно процентно съотношение GBR са с приоритет). Ресурсите се разпределят по равно между всички UE. Заявките, за които няма достатъчно ресурс, се отлагат във времето.



Фиг. 3. Блок-схема на Akyildiz приоритизация

Сравнителният анализ е реализиран между предложения механизъм за приоритизация в [5] със [7] и [8]. Сравнението е реализирано на базата на стойностите за Throughput, Packet Delivery Ratio и Delay за VoIP услуга за статични и мобилни UE и за non-GBR услуги за статични и мобилни абонати.

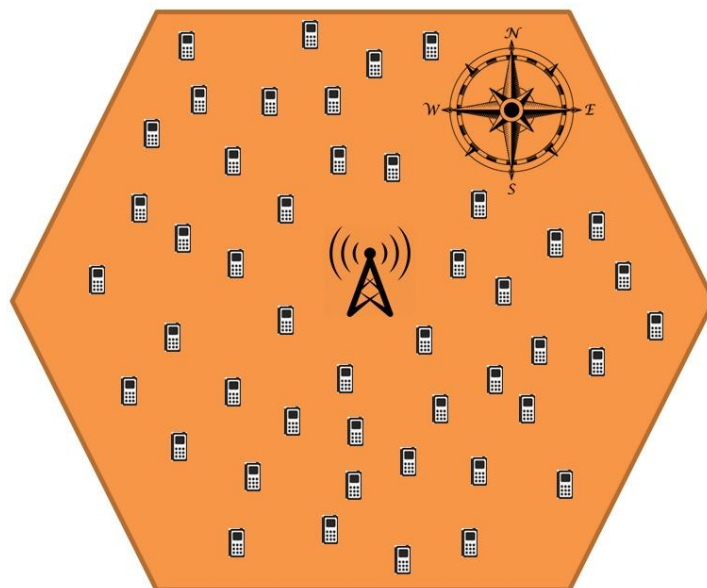
**Таблица 1.** Сравнителен анализ на разглежданите механизми за приоритизация

Приоритизационен алгоритъм Параметър	Предложен	QAPF	Akyildiz
Базиран на стандарта в [6]	Да	Частично	Не
Разпределение на ресурсите	Повече ресурс за UEs с по-висок приоритет	Равномерно разпределение на RB по абонати	Равномерно разпределение на RB по абонати
Параметри за приоритизация, различни от стандарта в [6]	Приоритет според заплащане за услуги; Разстояние до eNodeB; Скорост на движение на мобилни UE.	HOL Delay за GBR услуги	С по-висок приоритет са GBR или non-GBR в зависимост от процента на изискваните услуги
Отхвърлени заявки при обслужване	Не	Да	Не
Отложени заявки за обслужване в следващ таймслот	Да	Да	Да

### 3. Тестване и резултати от симулацията

Тестовите и получените резултати са реализирани на базата на предложения в [10] симулатор, който имплементира в себе си предложения, QAPF и Akyildiz механизми за приоритизация на трафика.

Опитната постановка е изградена на базата на топологията, представена на фигура 4. За всички тестове се използва LTE клетка с шест сектора, която работи с честотна лента от 20MHz и предоставя общо 100 PRB във всеки таймслот. За тестовите със статични абонати към eNodeB са свързани 100UE, а за тестовите с мобилни абонати към базовата станция са свързани 60 UE. Те са разположени на различно разстояние от eNodeB, като мобилните абонати се движат с различна скорост в произволна посока на движение. При тестовите статичните и мобилните абонати изискват VoIP или non-GBR услуга.



**Фиг. 4.** Топология на опитната постановка

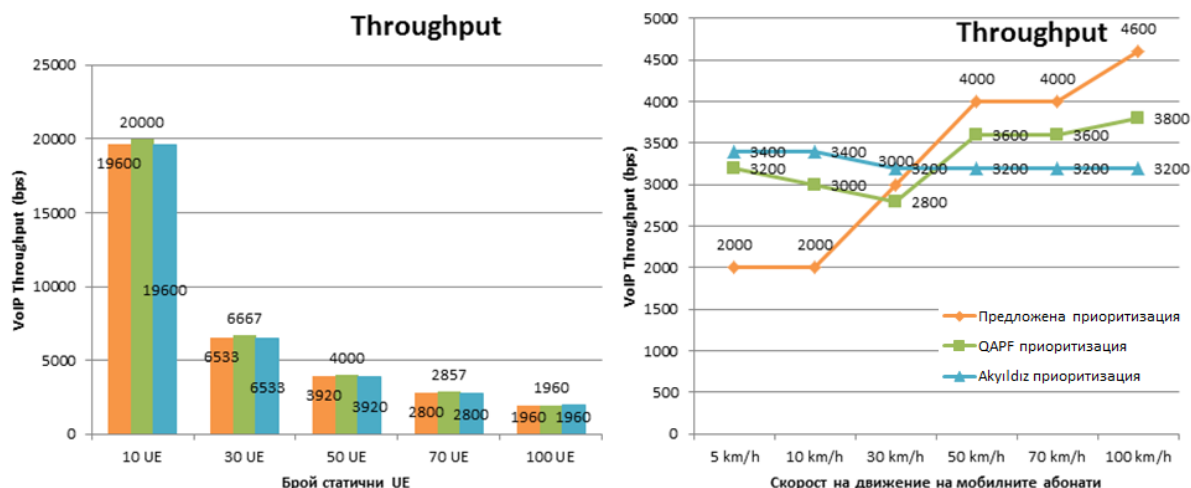
Направените тестове сравняват получените според трите механизма за приоритизация стойности за Throughput, Packet Delivery Ratio и Delay за VoIP и non-GBR услуга за статични и мобилни UE. Получените стойности са изчислени според уравнения (1), (2) и (3), като са представени средноаритметичните им стойности в рамките на един фрейм.

$$\text{Throughput} = \frac{\text{Number of RB sent in 1 Frame}}{\text{Transmission time}} \quad (1)$$

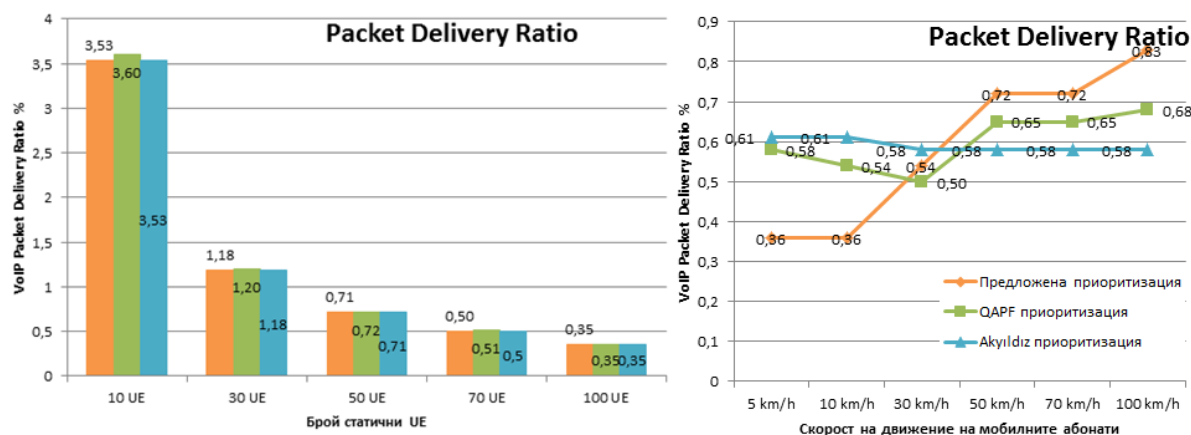
$$\text{PDR} = \frac{\text{Number of delivered RB in 1 Frame}}{\text{Sent RB}} * 100 \quad (2)$$

$$\text{Delay} = \frac{\text{Total time for deliver 1 Frame}}{\text{Total RB Sent}} \quad (3)$$

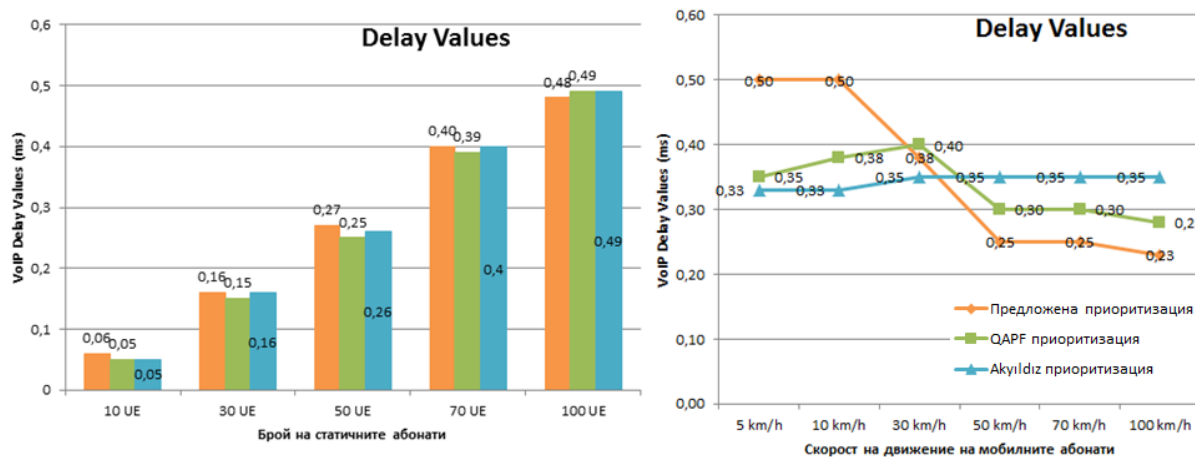
Според направените изследвания за VoIP услуга стойностите за Throughput и PDR и за трите механизма за приоритизация, при статични UE се променят обратнопропорционално с увеличаването на броя на абонатите. За мобилните абонати предложеният от автора и Myo механизми предоставят високи стойности за високоскоростните UE, а Akyildiz поддържа почти постоянни стойности, като предложеният алгоритъм позволява по-добри резултати (фигура 5 и фигура 6). Стойностите за Delay се променят правопропорционално с увеличаване на броя на статичните абонати. За мобилните абонати се вижда, че стойностите на забавянето се променят обратнопропорционално с увеличаване на скоростта на движение за предложения от автора алгоритъм и Myo, докато при Akyildiz са постоянни (Фиг. 7). За VoIP услугата получените стойности на изследваните параметри за статичните абонати са съизмерими за трите механизма за приоритизация, докато за мобилните UE предложеният алгоритъм предоставя по-добри резултати за високоскоростните абонати.



Фиг. 5. Throughput стойности за VoIP услуга при статични и мобилни UE

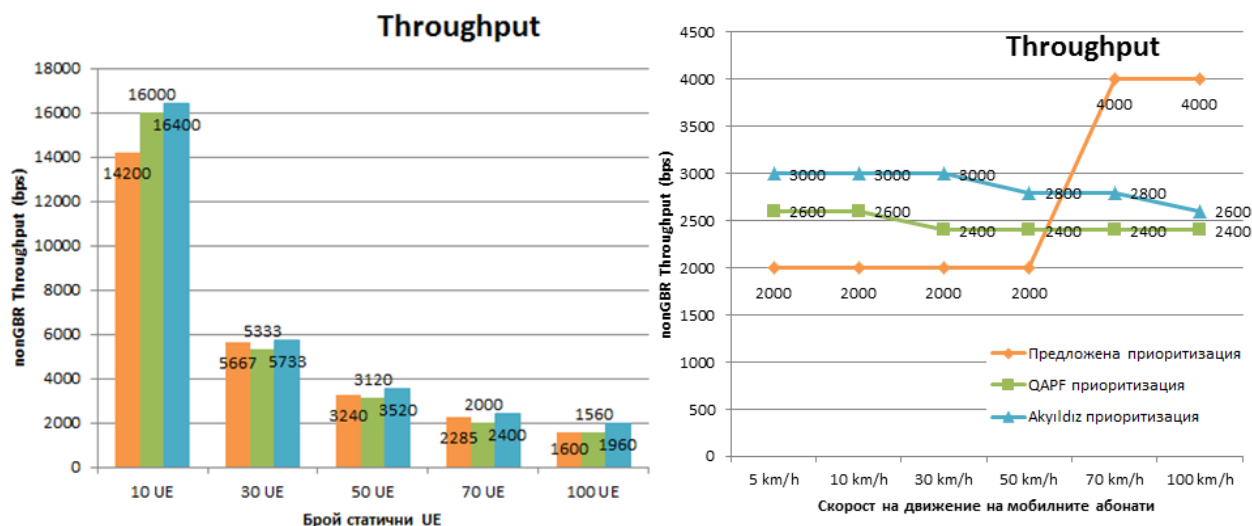


Фиг. 6. PDR стойности за VoIP услуга при статични и мобилни UE



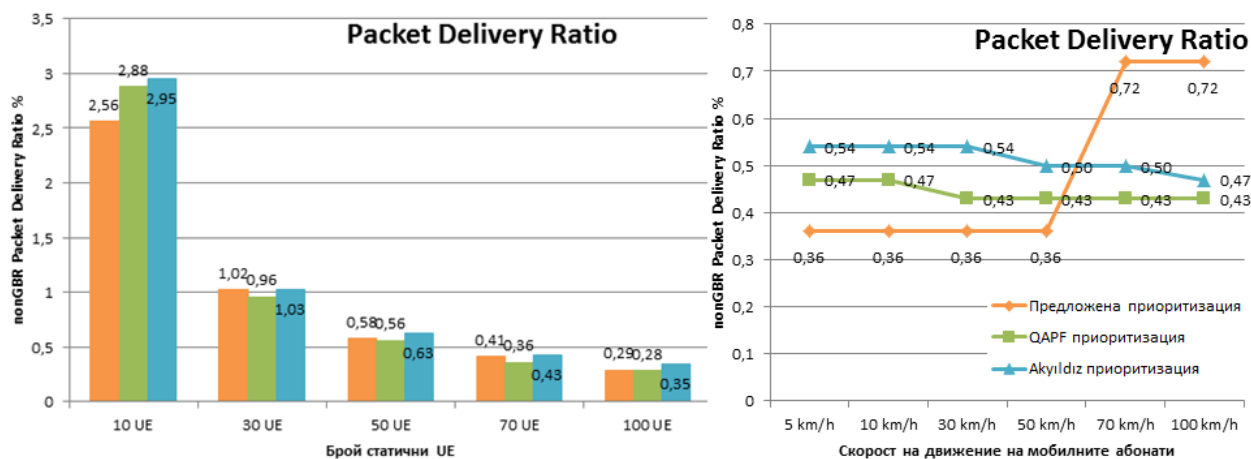
Фиг. 7. Delay стойности за VoIP услуга при статични и мобилни UE

Изследванията при изискани nonGBR услуги показват, че стойностите за Throughput и Packet Delivery Ratio се променят обратнопропорционално с увеличаване броя на UE и са съизмерими. За мобилните абонати обаче, предложеният механизъм позволява достигането на значително по-високи стойности за бързодвижещите се UE, докато при QAPF и Akyildiz приоритизацията те остават почти непроменени (фигура 8 и фигура. 9), като Akyildiz предоставя малко по-добри стойности от QAPF, защото при по-висок процент изискани non-GBR услуги ги обслужва с приоритет. Стойностите на Delay се променят правопрпорционално с увеличаване на броя на статичните абонати. При 100 UEs стойностите при предложения алгоритъм и на Муо са ниски, тъй като има отложени ресурси за обслужване в следващ фрейм, докато при Akyildiz отложените ресурси са по-малко, обслужените в рамките на фрейма абонати са повече, а стойностите на забавянето са по-високи. За мобилните абонати предложеният алгоритъм предоставя най-ниски стойности на забавянето за високоскоростните UE, докато при останалите не се променя, а QAPF и Akyildiz алгоритмите предоставят постоянни стойности на забавянето, независимо от скоростта на движение на абонатите (фигура. 10), защото разпределят ресурсите равномерно по абонати, без да отчитат скоростите на движение.

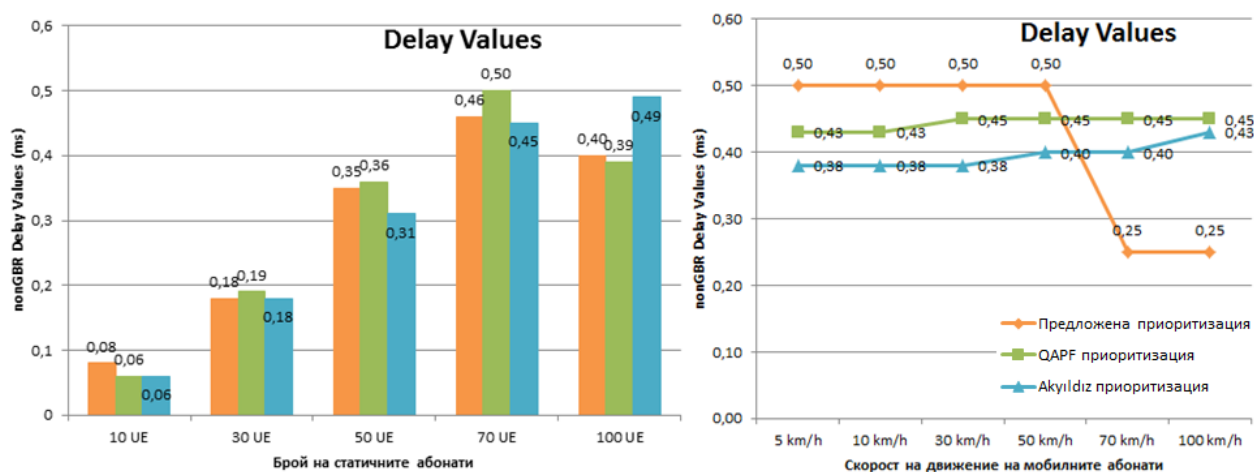


Фиг. 8. Throughput стойности за nonGBR услуга при статични и мобилни UE





Фиг. 9. PDR стойности за nonGBR услуга при статични и мобилни UE



Фиг. 10. Delay стойности за nonGBR услуга при статични и мобилни UE

#### 4. Заключение

При една и съща мрежа с един и същ брой UE използваният симулатор дава различни резултати за различните механизми за приоритизация. При еднакъв брой статични UE и за VoIP, и за nonGBR услуги с увеличаване на броя на абонатите Throughput и Packet Delivery Ratio на мрежата намаляват, а Delay се увеличава, защото с нарастване броя на потребителите ресурсите, необходими за тяхното обслужване, намаляват. При обслужване на статични абонати и трите алгоритъма дават почти съизмерими резултати, но QAPF и Akyildiz имат по-добри стойности от предложения механизъм, защото разпределят ресурсите по абонати равномерно, докато предложеният алгоритъм предоставя повече ресурс на потребителите с по-висок приоритет.

При мобилните абонати както за VoIP, така и за nonGBR услуги предложеният алгоритъм предоставя по-високи стойности за Throughput и PDR за бързодвижещите се UE и по-ниски стойности на Delay за тях. Това е така, защото при еднакви параметри за платен приоритет и разстояние до eNodeB за мобилни UE предложеният алгоритъм обслужва приоритетно бързодвижещите се абонати, така те получават повече ресурс от останалите. При QAPF алгоритъма стойностите за сравняваните параметри за мобилните абонати с увеличаване на скоростта са подобрени за VoIP услуга и постоянни за nonGBR услуга, защото с приоритет се обслужват GBR услугите, като тези заявки, които надвишават максималния бюджет за забавяне, се отхвърлят, а ресурсите за всички се разпределят равномерно, като за nonGBR при голям брой UE може да няма достатъчно ресурс за

обслужване. При Akyıldız механизма сравняваните параметри за мобилните абонати и при GBR, и при non-GBR услугите имат почти постоянни стойности, независимо от скоростта им на движение. Стойностите за GBR и non-GBR са такива, защото UE се приоритизират според процента на изискания тип услуга и ресурсите се разпределят равномерно по абонати. Така, ако GBR са с по-голям процент, тези заявки се обслужват с приоритет и обратно. Това позволява постоянни стойности за всички абонати, което в някои случаи може да не предостави достатъчно добро обслужване.

## Литература

- [1]. 4G LTE World Coverage Map – LTE, WiMAX, HSPA+, 3G, GSM Country List. 19 April 2018. <https://www.worldtimezone.com/4g.html>. Last visit on 27.07.2018
- [2]. LTE Overview. <http://www.3gpp.org/technologies/keywords-acronyms/98-lte>. Last visit on 27.08.2018
- [3]. Interim Ericsson Mobility Report. February 2018. <https://www.ericsson.com/en/mobility-report/reports/november-2017>. Last visit on 27.08.2018
- [4]. Mobile Technology Statistics – Global,
- [5]. <http://www.5gamericas.org/en/resources/statistics/statistics-global/>. Last visit on 27.08.2018
- [6]. Aleksieva V., Haka A., Modified Scheduler for Traffic Prioritization in LTE Network. //Proceedings of the Second International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'17), 2017, Volume 2, pp. 228-238
- [7]. 3GPP. Tech. Specif. Policy and charging control architecture, 3GPP TS 23.203 version 15.3.0 Release 15, (2018-06)
- [8]. MYO, Z. M., MON, M. T. QOS-AWARE PROPORTIONAL FAIR (QAPF) DOWNLINK SCHEDULER ALGORITHM FOR LTE NETWORK. //International Journal of Advanced Computational Engineering and Networking, July –2015, Volume-3, Issue – 7, pp: 19-23
- [9]. Akyıldız, H. A., Akkuzu, B., Hökelek, I., Çırpan, H. A. LTE Downlink Scheduler with Reconfigurable Traffic Prioritization. //IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom), Odessa, Ukraine, 27-30 May 2014
- [10]. Afroz, F., Heidery, R., Shehab, M., Sandrasegaran, K., Shompa, S. S. COMPARATIVE ANALYSIS OF DOWNLINK PACKET SCHEDULING ALGORITHMS IN 3GPP LTE NETWORKS. //International Journal of Wireless & Mobile Networks (IJWMN), October 2015, Vol. 7, № 5
- [11]. Aleksieva V., A. Haka. Simulation framework for realization of priority-based LTE Scheduler. //Journal Of The Technical University - Sofia, Plovdiv Branch, Bulgaria Fundamental Sciences and Application, 2017, Volume 23, pp. 101-105

**За контакти:**

ас. инж. Айдын М. Хъкъ  
катедра „Компютърни науки и технологии”  
Технически университет - Варна  
E-mail: aydin.mehmed@tu-varna.bg

# COMPARISON OF PERFORMANCES OF BUILT-IN TOOLS FOR LOAD AND EXPORT OF DATA TO AND FROM POPULAR DATABASES

Sinisa S. Ilic, Dragisa Miljkovic, Alempije Veljovic, Jasmina Novakovic  
Vladimir Veljovic

**Abstract:** There is often a need to transfer data from one source to another. Almost all modern databases have built-in tools that enable fast load and un-load of data to and from data tables or queries. In this paper, the comparison of performances of these tools is performed on popular databases (HP Vertica, MariaDB Columnstore, Oracle and MS SQL Server) and detailed description of their usage is given.

**Keywords:** database, data load, data export

## 1. Introduction

The amount of data that humankind store every day surpass order of several Quintillion ( $10^{18}$ ) bytes [1]. There is often a need to transfer data from one source to another; reason may be transferring from OLTP (that serve for transaction processing data) to OLAP (that serve for data analysis) databases, migration data from old to modern (new) databases or something else.

In process of transferring data from one to another database, usually ETL (Extract, Transform, Load) tools are used. It is not quite simple and intuitive to perform ETL process and use ETL tools. The ETL conceptual model should be developed first, not only to show an overview of the whole process, but also to map sources, targets and necessary transformations, verify transformation logic, ensure that requirements are met and the needed structures and content exist [2].

There is also a wide variety of such tools available in market – commercial and open source. In order to compare their different aspects of some popular ETL and data integration tools some authors [3-6] have analysed their advantages and disadvantages and dealt with user reviews on them.

But, sometimes it is just enough to export data from one database to, or import (load) data to another database from data files, not using online data transfer between databases (although the comparison of performances of loading data online and through data files is analysed [7]). It may be due to not existing physical connections or permissions for accessing databases or complicate use of existing ETL tools.

In this paper we firstly describe built-in commands and statements for batch data import and export in popular commercial and open source databases like: HP Vertica, Oracle DB, Microsoft SQL Server and MariaDB and then compared their performances experimentally. Our goal was to find out the order of time duration needed to load and un-load data to and from data tables and compare if results are of different order of magnitude. The databases used were prepared for OLAP.

## 2. Data used in experiment

For the experiments performed and described in the paper, the tables from Vmart scheme [8] are used. Vmart is data warehouse data mart used as a sample scheme in Vertica database virtual machine. This data mart consists of 11 dimensions and 4 fact tables. The names of data tables along with the number of attributes, n-tuples and other characteristics relevant for the goal of this paper for tables with sizes larger than 1 Mb are given in Table 1.

**Table 1.** Data from data tables used

Table name	Num. of columns	Num. of non-char columns	Number of records	Table size [Mb] (in CSV format)
online_sales_fact	18	17	5 000 000	361
store_sales_fact	15	13	5 000 000	358
store_orders_fact	19	18	300 000	31.3
inventory_fact	5	5	300 000	5.33
product_dimension	21	13	60 000	16.9
customer_dimension	21	10	50 000	6.53
employee_dimension	18	8	10 000	1.29

Some tables in Vmart scheme have varchar columns where a text of variable width can be found. The CSV format presented in Table 1 is Unix based (with linefeed – LF character as a row delimiter) and with coma (,) as a column delimiter.

### 3. Preparation of working environment for data export/import

The experiments were performed on VMware virtual machines – one per database and were ran on VMware Workstations 14 Player on a single Windows 10 host machine with Intel I7-6700HQ processor running on 2.60 GHz. The virtual machines were ran from internal sata3 HDD, and 6GB or RAM memory is allocated to each VM. Virtual Machines were not ran simultaneously. The VM with pre-installed HP Vertica was downloaded from HP Vertica web page [8] and other VMs were created from the scratch. The VMware tools were installed on every VM.

For each database, a SQL script with DDL commands (CREATE TABLE, ALTER TABLE) is prepared to create tables and relations of Vmart scheme. The initial vmart\_define\_schema.sql script can be found at pre-configured Vertica VM (in /opt/vertica/examples/VMart\_Schema/), but because of “dialects” of different DBs used, the SQL syntax must be changed accordingly for every specific database.

Today, data from OLTP are usually transferred to OLAP platforms by using modern ETL tools. In the ETL tools the processes consist of list of steps for Extracting, Transforming and Loading where once defined and stored steps remain mainly constant for a long time. One of typical steps in bulk data loading is temporarily disabling (or removing) data constraints (like: check, FK, PK) before loading and then after process (load) finishes, the constraints are implemented back again. For each database used in this study, the scrips for removing and establishing data constraints are prepared and implemented, so none of the measured time needed for data load is used to check data constraints.

**HP Vertica Analytic Database 9.0.1** Enterprise edition [9] is pre-installed on Centos 7 OS on Vmware virtual machine. The Vmart scheme is also installed and data in tables are loaded. The memory parameters are modified compared to the original installation: initial heap size is set to 2Gb and maximum heap size is set to 4Gb.

In order to export data from tables, the following vsql command is used from the standard linux shell:

```
vsql -U dbadmin -w password -c "SELECT * from schema.table_name" -Atq -F ',' -o "/home/dbadmin/table_name.csv"
```

In the vsql command presented -U is switch for username, -w for password, -c for command (SQL statement), -Atq disables command echo in CSV file (showing SQL statement on beginning of file), -F for field separator and -o for output filename.

The shown command does not measure the command time duration, so we used standard linux shell option – we added the following “timestamp” command:

```
echo $(( $(date +%s%N)/1000000 ))
```

that show system timestamp in milliseconds before and after vsql command using linux notation for separation of multiple commands in one line: “timestamp command && vsql command && timestamp command”. In this way the start timestamp is shown twice: immediately before and after execution of vsql command. The time duration of vsql command execution is then achieved by subtraction of two shown timestamps.

Import of data from CSV files is performed after truncating tables. It is done by using the “COPY” statement from built-in web user interface - Vertica Management Console:

```
COPY scheme.table_name FROM '/opt/vertica/examples/VMart_Schema/  
table_name.tbl' DELIMITER '|' NULL '' DIRECT;
```

As can be seen in the statement above, tables of vmart scheme can be loaded from already prepared .tbl files (csv files with “|” delimiters) in /opt/vertica/examples/VMart\_Schema/ directory. The duration of COPY statements running time are obtained using built-in Vertica (log) query table query\_requests by selecting fields: request and request\_duration (and by filtering with WHERE request like ‘COPY%’).

**MariaDB columnstore database** [10] version 10.2.13 is installed also on Centos 7 OS with some modifications in initial settings related to RAM memory size. In the file /usr/local/mariadb/columnstore/mysql/my.cnf the following parameters are set: key\_buffer\_size = 2048M, innodb\_buffer\_pool\_size = 4096M, innodb\_log\_file\_size = 1024M and in [isamchk] and [myisamchk] sections key and sort buffer sizes are set to 1024M. The MySQL Workbench application is used to access the database and to run SQL statements. For load of CSV files the following statement is used within the MySQL workbench:

```
load data local infile 'table_name.csv' into table scheme.table_name  
fields terminated by ',';
```

The duration of statement execution time is shown in MySQL Workbench for each statement.

Export of data from tables to CSV files is performed also in MySQL Workbench by using the following statement:

```
SELECT * FROM scheme.table_name INTO OUTFILE 'table_name.csv' FIELDS  
TERMINATED BY ',' LINES TERMINATED BY '\n';
```

**Oracle DB Server** 12c release 12.1.0 Enterprise [11] is installed also on Centos OS. After installation of Oracle DB Server, the following settings are modified: memory\_target is set to 4Gb, sga\_max\_size and sga\_target to 2Gb and inmemory\_size parameter is set to 1Gb. During initial installation the configuration with “pluggable” databases is selected and inmemory\_size is set also to 1Gb for pluggable database. All tables from Vmart scheme are created in pluggable database. For connection to Oracle DB, its standard development tool “SQL Developer” is used. To import data from CSV files the Sql Loader command is used from linux command shell using oracle user:

```
sqlldr userid=orauser@PDB/orapass, control=table_name.ctl,  
data=path/table_name.csv, direct=TRUE
```

where orauser is user defined in Oracle DB with permissions to access and insert data into table, PDB is connection name for pluggable database, orapass is password for Oracle user,

table\_name.ctl is name of control file where 'log' filename, 'bad' filename (with list of bad records), destination Oracle scheme, table\_name and format of columns are stated. This file contains 'instructions' for SQL Loader how to read and load data from data file. The format of control file can be found easy on the Internet.

The time duration needed for SQL Loader to load table is measured as in Vertica DB case by using Linux command to show timestamps before and after execution of SQL Loader command.

We tried several scenarios of exporting data to CSV files: by using built-in export database option from SQL Developer and by running script file (we can call it here: table\_name\_script.sql) from SQL Plus editor. Using export database option in SQL Developer is very intuitive to use, but we couldn't measure the time duration of data export; we could measure only time needed for Oracle DB to send data from a table to SQL Developer (from column elapsed\_time in system view: v\$sql), but not the time needed to write pulled data to a file. So, we created the script file:

```
SELECT systimestamp from dual; -- show system time

SET heading off

set pagesize 0    -- remove header

set termout off  -- don't show query results

spool /home/oracle/inFact.csv -- open file

select x from (

select col1||','||col2||','||col3||','||col4||','||colN as x

from scheme.table_name);

spool off; -- close file

set termout on -- show query results

SELECT systimestamp from dual; -- show system time
```

and ran it from SQL Plus editor with the following command: @path/table\_name\_script.sql. Command "SELECT systimestamp from dual;" from the script is used to print a timestamp in SQL Plus editor before and after spooling (sending a query results to a file) in order to be measured execution time of exporting data to csv file. The nested subquery here helps in cutting trailing blank spaces at the end of rows that would appear if only "Select col1||','||colN from scheme.table\_name" would exist.

**MS SQL 2017 Server** (ver. 14.0) Standard edition [12] is installed on MS Windows Server 2016. The default settings for SQL Server related to memory size are remained (maximum memory set to 2,147,483,647Mb). The SQL Management studio is used to load data into SQL Server tables in two options: *BULK INSERT* statement and *INSERT INTO SELECT \* from OPENROWSET()*, and BCP utility is used to load data from command prompt.

In order to measure time duration of data load, two datetime variables are declared: @time\_start and @time\_end and values from system function getdate() are used to assign system times to these variables before and after execution of INSERT statements. The function datediff is then used to calculate the time difference. The *BULK INSERT* statement is formulated as:

```
DECLARE @bulk_cmd varchar(1000);

SET @bulk_cmd = 'BULK INSERT scheme.table_name FROM ''table_name.csv''

WITH (fieldterminator = ',',',', ROWTERMINATOR = '''+CHAR(10)+'')';

EXEC(@bulk_cmd);
```



Char(10) denotes Linefeed (0x0A) – unix based end of line. Another statement has the format:

```
INSERT INTO scheme.table_name  
SELECT * FROM OPENROWSET (BULK 'table_name.csv', FORMATFILE =  
'table_name_bcp.fmt') as T
```

File table\_name\_bcp.fmt is format file and has description of number of fields in csv file, column names and column types, field and row separators, collation types etc. The description of format file can be found on the Internet. The same file is used also for BCP utility that can be run from a command prompt:

```
bcp scheme.table_name IN "table_name.csv" -f "table_name_bcp.fmt" -e  
"error.log" -T -b 5000000 -a 65535
```

The key word for data load is IN, -e is switch for the error log file name, -T is switch for Trusted connection (bcp can connect to the database installed on the same computer without using username and password), -b is a switch for batch size (maximum number of rows to be loaded at once) and -a is a switch for a packet size that can be from 4096 to 65535 bytes. BCP utility shows the elapsed time after successful execution. For last two methods we had to replace LF character with LF CR characters in CSV files using unix2dos command.

To export data from MS SQL server to a CSV file we used only BCP utility from a command prompt with the following syntax:

```
bcp scheme.table_name OUT "table_name.csv" -T -c -t, > export.log
```

One can see that the key word here is OUT for exporting data, -c switch to export data using a character data type (.csv is text based file), -t is switch for field delimiter and ">" redirects log from screen to log file where elapsed time of execution can be found.

#### 4. Results

In Table 2 the time durations needed for load of data from data files to data tables are presented. It can be seen that Oracle SQL Load tool has the best performances. It loads tables: online\_sales\_fact and store\_sales\_fact with 5 million rows and with more than dozen columns in 13.5 and 12.5 seconds respectively. The second best result is for HP Vertica COPY statement with durations of 26.5 and 18.3 seconds. MariaDB columnstore "load data" statement imports these tables for 46.5 and 48.7 seconds and BCP utility loads them into MS SQL for 42.9 and 63.7 seconds. If we compare the bulk load durations of BCP utility, "bulk insert" and "insert into select \* from openrowset" statements, we may see that they are pretty same with slight differences in favor of the BCP. But, all utilities load data of 5 million rows for less than a minute.

**Table 2.** Duration of data load from data files to database tables in seconds

Table name	Vertica	MariaDB	Oracle	MS SQL Server		
				BULK Insert	Open Rowset	BCP
online_sales_fact	26.573	46.580	13.455	55.767	78.993	42.906
store_sales_fact	18.321	48.712	12.561	68.140	63.740	63.750
store_orders_fact	1.765	28.341	0.968	15.277	11.390	3.797
inventory_fact	0.496	6.258	0.696	1.114	1.816	1.485
product_dimension	0.313	4.251	0.887	5.203	5.080	3.422
customer_dimension	0.254	4.469	1.214	1.677	1.284	0.438
employee_dimension	0.096	3.230	0.335	0.140	0.140	0.156

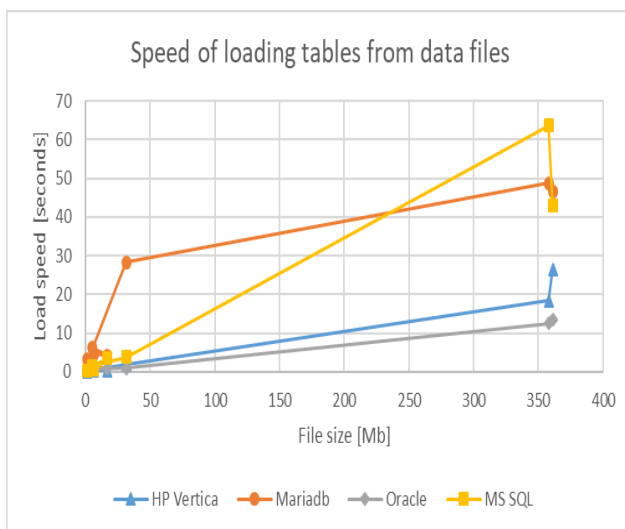


The results are different for exporting data from database tables to files and performances are shown in Table 3. HP Vertica's "vsql -o" utility exports aforementioned tables for 13.5 and 14.2 seconds, MariaDB "select \* from into outfile" statement exports for 18 and 15.4 seconds while SQL BCP utility and Oracle spool statement export data for 48.5, 28.5 and 113.4 and 64.1 seconds.

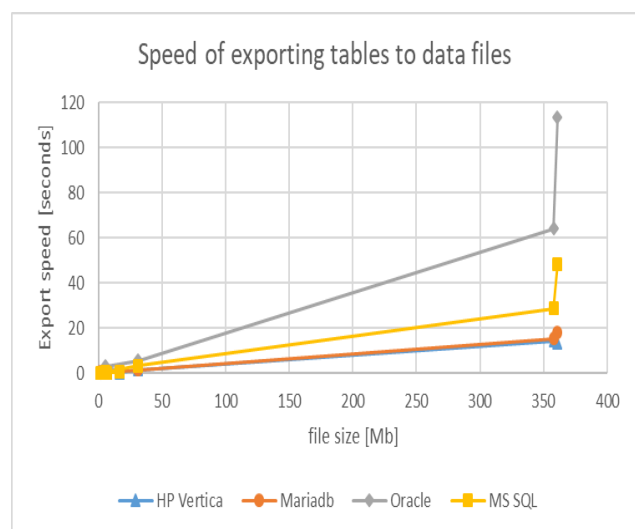
**Table 3.** Duration of exporting data from database tables to data files in seconds

Table name	Vertica	MariaDB	Oracle	MS SQL
online_sales_fact	13.553	17.993	113.389	48.547
store_sales_fact	14.239	15.375	64.148	28.594
store_orders_fact	1.446	1.624	5.506	3.203
inventory_fact	0.031	0.370	3.174	0.453
product_dimension	0.211	0.969	1.061	0.438
customer_dimension	0.231	0.924	0.850	0.281
employee_dimension	0.070	0.637	0.199	0.063

The visual presentations of times needed to load and export tables of different sizes are shown in figures Fig. 1 and Fig. 2. The BCP utility performances are used for MS SQL in both figures.



**Fig. 1.** Duration of data load



**Fig. 2.** Duration of data export

## 5. Conclusions

The transfer of data between different information systems is usually performed through ETL tools that can perform extract, transform and load operations online. But, sometimes ad-hoc transfer of data can be performed through data files. That is why all commercial and open source databases have built-in tools for loading and exporting batch data from/to data files.

From performed experiments it can be seen that databases built-in tools can load into tables and un-load from tables a million of rows very fast from / to data files. There are slight differences between the speeds of tools of different databases, but not for the order of magnitude. For the largest tables in the experiment, that contain 5 million of rows and more than dozens of columns, the load time was less than a minute, while export time was less than two minutes in a worse case.

The best load time for a large table is achieved by Oracle SQL Loader and the best export time is achieved by HP Vertica "copy" statement.

## Acknowledgment

This work has been supported by the Ministry of Science and Technological Development of the Republic of Serbia under Project No.TR-35026.

## References

- [1]. Every Day Big Data Statistics, <http://www.vcloudnews.com/every-day-big-data-statistics-2-5-quintillion-bytes-of-data-created-daily/>
- [2]. Saša Dupor and Vladan Jovanović, An approach to conceptual modelling of ETL processes, Proceedings of International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014, Opatija, Croatia
- [3]. Nilesh Mali and SachinBojewar, A Survey of ETL Tools, International Journal of Computer Techniques, Volume 2 Issue 5, Sep – Oct 2015
- [4]. Rajendrani Mukherjee and Pragma Kar, A Comparative Review of Data Warehousing ETL Tools with New Trends and Industry Insight, Proceedings of the 2017 IEEE 7th International Advance Computing Conference (IACC), Hyderabad, Telangana, India
- [5]. J.P.A. Runtuwene et al., A Comparative Analysis of Extract, Transformation and Loading (ETL) Process, ICIEVE 2017, IOP Conf. Series: Materials Science and Engineering, 306, (2018) 012066 doi:10.1088/1757-899X/306/1/012066
- [6]. Comparing Microsoft, Oracle, <https://www.gartner.com/reviews/market/data-integration-tools/compare/microsoft-vs-oracle>
- [7]. Just How Fast is ODBC? A “Loaded” Comparison. <http://www.iri.com/blog/data-transformation2/sql-loader-odbc-comparison/>
- [8]. Vertica Downloads, Community Edition  
<https://my.vertica.com/download/vertica/community-edition/>
- [9]. Vertica Documentation, Vertica Analytic Database  
[https://my.vertica.com/docs/9.0.x/PDF/Vertica\\_9.0.x\\_Complete\\_Documentation.pdf](https://my.vertica.com/docs/9.0.x/PDF/Vertica_9.0.x_Complete_Documentation.pdf)
- [10]. MariaDB ColumnStore <https://mariadb.com/kb/en/library/mariadb-columnstore/>
- [11]. Oracle Database Online Documentation 12c Release 1 (12.1)  
<https://docs.oracle.com/database/121/>
- [12]. SQL Server Documentation <https://docs.microsoft.com/en-us/sql/sql-server/sql-server-technical-documentation?view=sql-server-2017>

### For contacts:

Dr Sinisa Ilic

E-mail: [sinisa.ilic@pr.ac.rs](mailto:sinisa.ilic@pr.ac.rs)

Dragisa Miljkovic

E-mail: [dragisa.miljkovic@pr.ac.rs](mailto:dragisa.miljkovic@pr.ac.rs)

Department of Computer science and Informatics

University of Pristina in Kosovska Mitrovica, Serbia

Alempije Veljovic

E-mail: [alempije@beotel.net](mailto:alempije@beotel.net)

Vladimir Veljovic

E-mail: [veljo09@gmail.com](mailto:veljo09@gmail.com)

Faculty of Technical Sciences Cacak, University of Kragujevac,

Kragujevac, Serbia

Jasmina Novakovic

E-mail: [jnovakovic@sbb.rs](mailto:jnovakovic@sbb.rs)

Belgrade Business School, Higher Education Institution for Applied Studies,

Belgrade, Serbia

# ДЕСКТОП ПРИЛОЖЕНИЕ ЗА ОБУЧЕНИЕ В КОДИРАНЕ НА ДАННИ И ВИЗУАЛИЗАЦИЯ НА МОДУЛАЦИИ

Венета П. Алексиева, Владислав Г. Димитров

**Резюме:** Предложеното настолно приложение реализира най-популярните алгоритми за компресиране на данни без загуба. То визуализира кодираните низове с основните модуляции. Идеята е да се представи един и същ текст на съобщение в друга кодова схема и да се види как той изглежда като модулиран низ в съответната преносна среда. Това приложение е изготвено с цел да се използва за визуализации в упражнения, свързани с кодиране в преносните среди при мрежови комуникации.

**Ключови думи:** Аналогови и цифрови модуляции, компресия на данни, кодиране по Шенон-Фано, кодиране по Хафмън.

## Desktop application for education in data encoding and vizualization of modulations

Veneta P. Aleksieva, Vladislav G. Dimitrov

**Abstract:** The proposed desktop application realizes most popular lossless data compression algorithms. In addition, it visualizes coded strings with the basic modulations. The idea is to present the same message text in different code scheme and see how they looks as modulated string in the network media. This application is completely useful in labs on subjects, related to network communications.

**Keywords:** Analog and Digital Modulations, Data Compression, Shannon-fano coding, Huffman coding.

## 1. Въведение

Експерименталният опит има значителен ефект върху инженерното обучение на студентите и подобрява техните практически умения. Експериментите обаче, са с високи разходи, изискващи лаборатории със скъпо оборудване, което в усвояването на начални знания не се ползва в пълния му капацитет. Именно в подобни ситуации са подходящи десктоп приложения, които са икономически изгодни решения, тъй като не изискват нито лаборатории, нито експериментално специализирано оборудване.

В съвременните комуникационни технологии се прилагат усъвършенствани методи за пренасяне на спектъра на сигнала-носител на информация във високочестотната област, чрез изменение на параметрите на високочестотен сигнал, под действието на сигнала-носител на информация. Например в 2G клетъчни системи като GSM, DECT стандарт за преносими телефони, спътникови системи, PON мрежи се прилага Time division multiple access (TDMA), в HIPERLAN/2 IEEE 802.16a WiMax, Bluetooth се прилага Dynamic TDMA, за сателитна комуникация се прилага FDMA, CDMA се използва в 3G на GSM, а OFDMA се прилага в 802.11a,g,n,ac, HiperLAN/2, Mobile WiMAX, LTE, цифрово радио и цифрова TV. От друга страна, вероятността за възникване на грешки в комуникационните канали на съвременните конвергиращи мрежи е от порядъка на  $1.10^{-24}$  и за да бъде наблюдавана грешка, единственото решение е да се симулира подобна ситуация. Затова при използване на съвременно оборудване за демонстрация не могат да бъдат пряко наблюдавани нито основните цифрови модуляции в чист вид, нито методи за кодиране на данни, целящи тяхната компресия.

За постигане на максимално добри резултати с минимално време се търсят персонализирани компютъризирани решения за обучение чрез визуализация.

## 2. Анализ на съществуващите решения

Един от най-добрите подходи за обучение по компютърни комуникации е използването на специализирано оборудване за демонстрации - за аналогови модуляции [1] и за цифрови модуляции [2]. Икономически неприложимо е подобно решение за базови дисциплини, в които това оборудване ще се използва само в 1-2 лабораторни упражнения.

Друг по-икономичен подход към този момент в обучението по компютърни комуникации е използването на симулационна среда. Най-реалистична симулация на техники за аналогови и цифрови модуляции се реализира с помощта на MATLAB и Simulink [3]. За симулация и визуализация на по-сложни модулационни схеми, както и за изследване на съвременни комуникационни технологии се ползва основно Simulink [4,5]. На тази база от реална лаборатория се преминава към виртуална лаборатория, която може да бъде предоставяна дистанционно на студентите за самоподготовка и затвърждаване на придобитите знания. Възниква обаче необходимостта от подготвяне на обучителен курс с упражнения, подобен на [6], където под формата на учебник са представени актуални решения за обучение в комуникационните системи, базирано на примери, използващи MATLAB и Simulink, а в [7] обучението за цифрова обработка на сигнали се базира на примери, реализирани с MATLAB. Подобни решения обаче са подходящи за самообучение и не са персонализирани към конкретен курс или целева група. Подобен подход изисква от студентите познаване на тези симулационни среди, което предполага в осъвременените учебни планове да фигурира дисциплина с подобна насоченост, предхождаща дисциплините, свързани с компютърни комуникации.

В [8] се предлага уеб-базиран образователен интерфейс AnalogWeb, който е разработен с помощта на MATLAB Builder за NE с WebFigures и ASP.NET платформа. Той не изисква познаване на средата MATLAB и осигурява съвременни инструменти за визуализация като увеличение, движеща се графика и триизмерно изображение, но само за аналогови модуляции, което в обучението по компютърни комуникации е недостатъчно.

Настоящият доклад представя десктоп приложение за виртуална лаборатория, предназначено за обучение на студенти в основи на компютърните комуникации. Чрез него се осъществява визуализация на базови модуляции и студентите се обучават на различни методи за кодиране на данни.

## 3. Проектиране на приложението

За десктоп приложение за кодиране на данни и визуализация на модуляции е избрана среда Microsoft Visual Studio 2012 и език за програмиране C#.

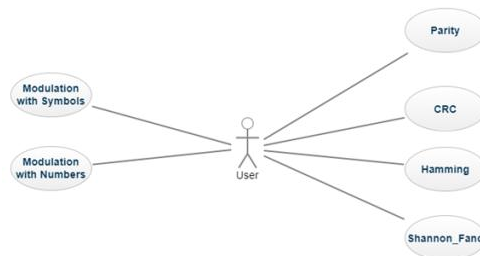
Функционалните изисквания към приложението са:

- да позволява на потребителя да въвежда своята информация (поредица от ASCII символи или числа) и на база на нея да се визуализира избраната модулация - амплитудна, честотна, фазова;
- да позволява на потребителя да въвежда своята информация и на база на нея да демонстрира нейното кодиране по различни алгоритми - код по четност, цикличен код, код на Хеминг и код на Шенон-Фано;
- да позволява на потребителя да проверява и открива с приложението грешки на вече кодирана информация.

Нефункционалните изисквания са:

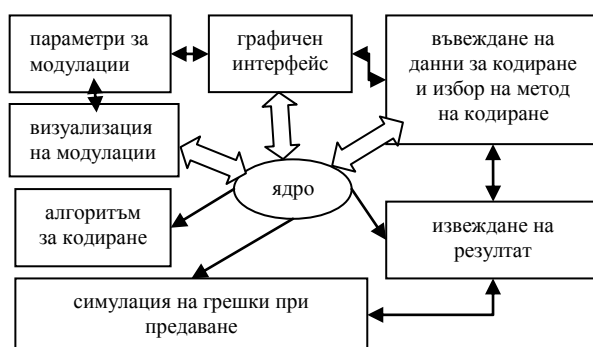
- интерфейсът на приложението да е на английски език, за да се даде възможност да се използва в курса на обучение и на български, и на английски език;
- опростен и интуитивен интерфейс;
- автоматична визуализация на модулациите и резултатите от кодирането.

На фигура 1 е представена use-case диаграма на приложението. Възможностите за извеждане на амплитудна, честотна и фазова модулация са както за въведени символи (техния ASCII код), така и за числа, преобразувани в бинарен код. От друга страна въведените данни могат и да се кодират с код по четност, цикличен код, код на Хеминг и код на Шенон-Фано.



Фиг. 1. Use-case диаграма

С цел универсалност на създаденото приложение и поради факта, че се прилага в обучение както на български, така и на английски език, е избран английски език. Архитектурата на приложението е представена на фигура 2.

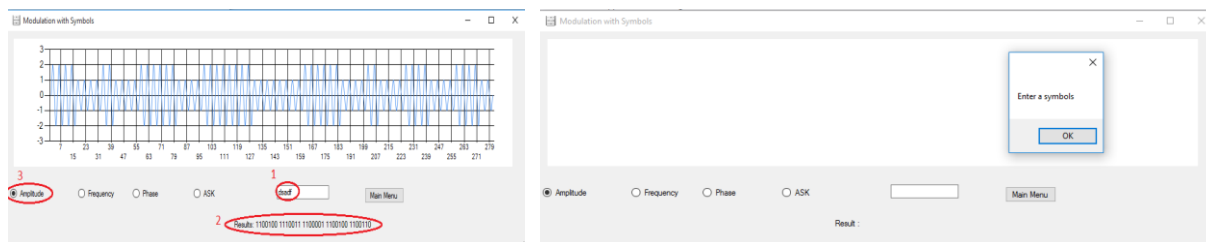


Фиг. 2. Архитектура на приложението

С оглед на целевата група, за която е предназначено приложението и съгласно заложеното в курса по основи на комуникациите, където се ползва, са избрани стандартни алгоритми за реализация, като за реализация са предвидени основните видове модулации - амплитудна, честотна и фазова модулация, като фазовата модулация е с фаза  $180^0$ . За реализация на кодиране на символи, представени с ASCII код в двоичен вид и на числа, преобразувани в двоична бройна система, са избрани код по четност, матричен код по четност, цикличен код с 5 различни полинома, код на Хеминг и код на Шенон- Фано.

#### 4. Реализация

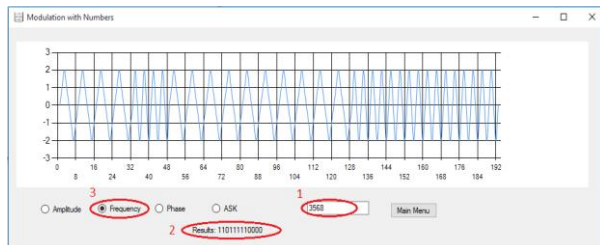
На фигура 3 (вляво) е представена визуализация на модулации с всякакви символи като входни данни.



Фиг. 3. Визуализация на модулации на символи и съобщение за грешки

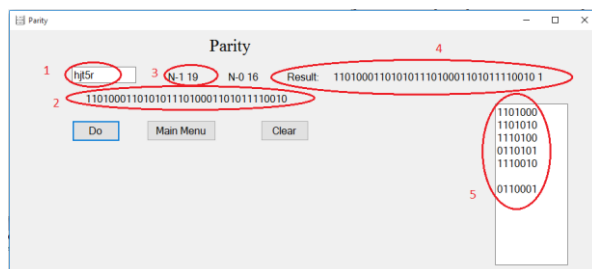
Поле 1 е за въвежданата от потребителя поредица от символи. Поле 2 е за представянето на въведените символи в двоична последователност, получена от ASCII кодовете на символите. Поле 3 е за избор на модулация (на фигурата е избрана амплитудна модулация) от потребителя. Ако не се въведат символи, се визуализира показаното съобщение на фигура 3 (вдясно).

На фигура 4 е представена визуализация на модуляции с число, което се преобразува в двоичен код. Поле 1 е за въвежданото от потребителя число. Поле 2 е за представянето на въведеното число в двоично. Поле 3 е за избор на модулация (на фигурата е избрана честотна модулация) от потребителя. Ако не се въведат символи, се визуализира съобщение, подобно на показаното на фигура 3.



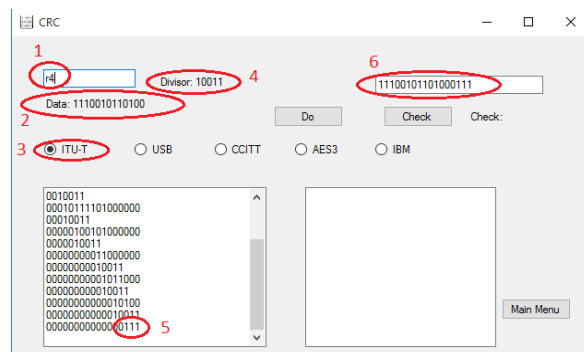
Фиг. 4. Визуализация на модуляции на числа

На фигура 5 е представено шумоустойчиво кодиране с контрол по четност. Потребителят въвежда последователност от символи в полето 1. Визуализира се двоичният им вид в поле 2 и се проверява броят на 1-ци (поле 3) и ако той е нечетен брой, в края на двоичния низ се добавя 1-ца, в противен случай се добавя 0. В поле 4 се визуализира резултатът след контрола по четност. В същия прозорец се реализира и матричен код по четност, където последователността от разряди се разделя на групи. Групите се записват в матричен вид и се проверяват по четност стълбовете на получената матрица. Резултатът се визуализира в поле 5. Ако не се въведат символи, се визуализира съобщение за грешка.



Фиг. 5. Контрол по четност

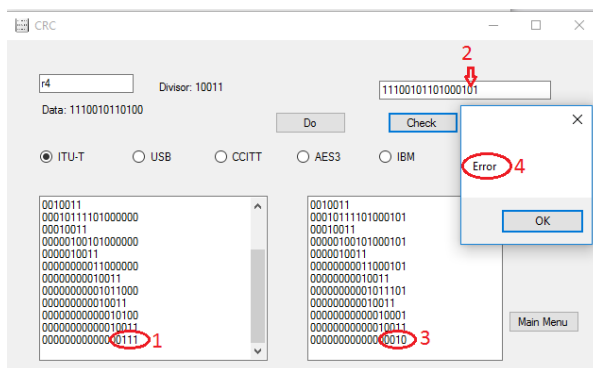
На фигура 6 е представена реализацията на цикличен код (CRC). В поле 1 потребителят въвежда поредица от символи.



Фиг. 6. Цикличен код

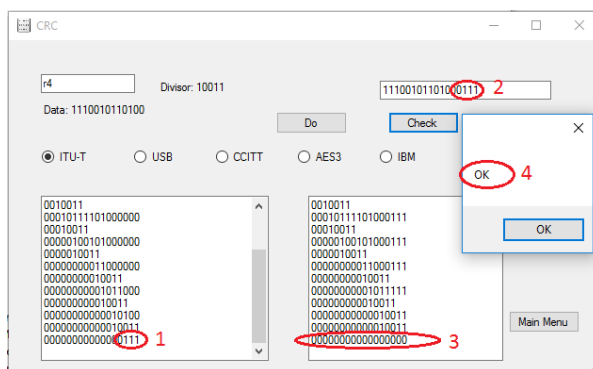
В поле 2 се визуализира двоичният вид на поредицата от символи. В поле 3 се избира с кой полином да бъде кодирана информацията, който се визуализира в поле 4. В поле 5 се представя резултатът от деленето и визуализация на кодиращите битове. В поле 6 се представя долепянето на кодиращите битове към основната информация.

На фигура 7 е представена проверка на вече кодираната информация с CRC, като в случай, че информацията не е коректна, се извежда съобщение за грешка (поле 4). В поле 1 е представен резултат от деленето и визуализация на кодиращите битове. В поле 2 се дава възможност потребителят да промени стойността на точно 1 от битовете. При избор на възможността за проверка (бутон "Check") в поле 3 се извежда резултатът. В този случай се вижда, че при грешка не се получават само нули, което е и признак, че има грешка при предаването на информацията.



Фиг. 7. Симулация на грешка при предаване с CRC

На фигура 8 е представена проверка на вече кодирана информация по CRC, в случай, че информацията е коректна и няма грешка при предаването на информация (поле 4). В поле 1 е представен резултат от деленето и визуализация на кодиращите битове. В поле 2 се вижда как битовете автоматично се добавят накрая на информацията. В поле 3 се вижда, че след проверка на информацията и като резултат има само 0 (информацията е предадена без грешки).



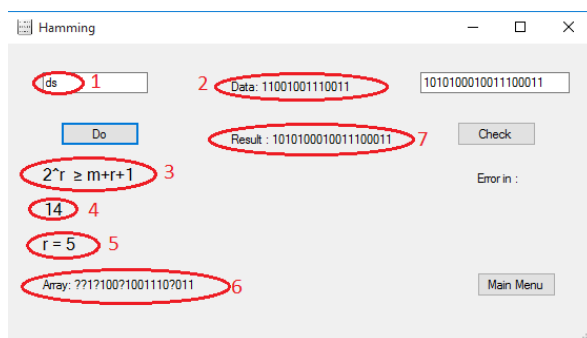
Фиг. 8. Кодирание с CRC и предаване без грешки

На фигура 9 е представен код на Хеминг. В поле 1 се въвеждат от потребителя символи. В полета 2, 3, 4, 5 се визуализират съответно двоичният вид на въведените от потребителя символи, формулата, по която се изчислява броя на кодиращите битове, дължина на преобразувания в двоичен вид символен низ, необходимият брой кодиращи битове. В поле 6 е представен низът преди прилагане на алгоритъма, с който се изчисляват кодиращите битове, а в поле 7 е визуализиран вече кодираният низ.

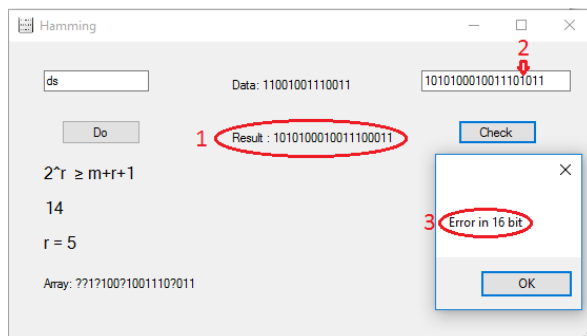
На фигура 10 е представена проверка на вече кодираната по код на Хеминг информация дали има открита грешка. В поле 1 е визуализирана вече кодираната информация, в поле 2 е дадена възможност потребителят да промени бит в кодираната информация, симулирайки



грешка в процеса на предаване, а в поле 3 се представя в кой бит е грешката след прилагане на алгоритъма за проверка.

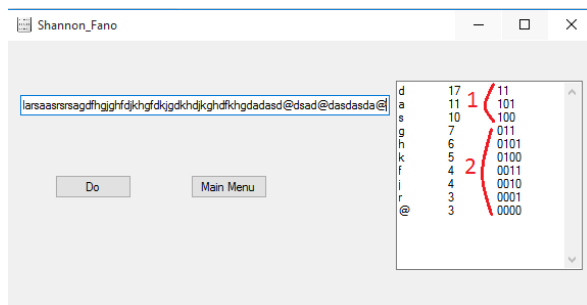


Фиг. 9. Кодиране с код на Хеминг



Фиг. 10. Откриване на грешка с код на Хеминг

На фигура 11 е представен код на Шенон–Фано. Потребителят въвежда последователност от символи. Ако няма въведени символи, се визуализира подканващо съобщение да се въведат такива. Ако има въведени символи, се създава таблица с вероятността за тяхното срещане. На тази база се получават кодовете на всеки от символите.



Фиг. 11. Код на Шенон-Фано

## 5. Заключение

В доклада е предложено десктоп приложение за виртуална лаборатория, предназначено за обучение на студенти в основи на компютърните комуникации. Чрез него се осъществява визуализация на базови модуляции и студентите се обучават на различни методи за кодиране на данни. Приложението позволява да се симулира възникване на грешки при предаване в комуникационните канали, за да се демонстрира работата на съответната схема за кодиране. Към момента се прилага в дисциплината "Основи на компютърните комуникации" в ОКС "Бакалавър" на специалност "Компютърни системи и технологии" както в обучението на български език, така и в обучението на английски език. За бъдещо развитие модулната структура на приложението позволява да бъде разширено с други кодове и модуляции.

## Литература

- [1]. XK-GP1 Analog communication training kit for college, electronic training kit, [https://www.alibaba.com/product-detail/XK-GP1-Analog-communication-training-kit\\_60514542723.html](https://www.alibaba.com/product-detail/XK-GP1-Analog-communication-training-kit_60514542723.html) [last visited on 2018]
- [2]. Communication Theory Experiment Trainer, [https://www.alibaba.com/product-detail/Communication-Theory-Experiment-Trainer\\_216108182.html](https://www.alibaba.com/product-detail/Communication-Theory-Experiment-Trainer_216108182.html) [last visited on 2018]
- [3]. MathWorks, Simulink, <https://www.mathworks.com/products/simulink.html> [last visited on 2018]
- [4]. Mr.P.Sunil Kumar, Dr.M.G.Sumithra, Ms.M.Sarumathi, Performance evaluation of Rayleigh and Rician Fading Channels using M-DPSK Modulation Scheme in Simulink Environment //IJERA, ISSN: 2248-9622, Vol. 3, Issue 3, May-Jun 2013, pp.1324-1330
- [5]. Mendis G., J. Wei, A. Madanayake, Deep learning-based automated modulation classification for cognitive radio, IEEE International Conference on Communication Systems (ICCS), 14-16 Dec. 2016.
- [6]. Choi K., H. Liu, Problem-Based Learning in Communication Systems Using MATLAB and Simulink, p.400, ISBN: 978-1-119-06034-5, 2016.
- [7]. Giron-Sierra J. M., Digital Signal Processing with Matlab Examples, Volume 1: Signals and Communication Technology, p.533, DOI 10.1007/978-981-10-2534-1, Springer Science, 2017.
- [8]. Kaçar S., C. Bayılmış, A Web-Based Educational Interface for an Analog Communication Course Based on MATLAB Builder NE With WebFigures, IEEE Transactions on Education, Volume: 56, Issue: 3, pp. 346 - 354, Aug. 2013.
- Baumann, W., B. Herberg. Papierchemikalien, Berlin, Springer-Verlag, 1994, pp. 305.

### За контакти:

доцент д-р инж. Венета П. Алексиева  
катедра „Компютърни науки и технологии”  
Технически университет-Варна  
E-mail: VAleksieva@tu-varna.bg



# DESIGN AND IMPLEMENTATION OF WEB APPLICATION FOR ANALYSIS AND COMPARISON OF HYPERTEXT TRANSFER PROTOCOLS SPDY AND HTTP/2

Vladimir Dimitrov, Georgi Gyurov

**Abstract:** This article examines the design and the implementation of progressive web application which purpose is for analysis and comparison of protocols for hypertext transfer – SPDY and HTTP/2. In 2015 Internet Engineering Task Force (IETF) the HTTP working group developed the second major version of the protocol in the form of HTTP/2. In May 2015, the HTTP/2 specification is officially standardized in response to the SPDY protocol from Google.

**Keywords:** HTTP/2, SPDY, HTTP, Service Workers, Single Page Applications (SPA), AngularJS

## 1. Introduction

The aim of this work is the design and the implementation of a software system (progressive web application) which purpose is for analysis and comparison of two protocols used for transfer of a hypertext – SPDY and HTTP/2. For achieving this aim the work is consisting of a two main parts:

1. Developing of a progressive web application using AngularJS as a framework, offline support through Service Workers for providing the best user experience and using the recent CSS code writing trends.
2. Analysis and comparison of SPDY and HTTP/2 by deploying the system onto server, supporting HTTP/2, monitoring and gathering of a key metrics, which are used in the final comparison of the two protocols.

## 2. Designing the architecture of the testing system

The basic principle followed in the design of the system is related to the idea for an application, developed on a single web page, namely Single Page Applications (SPA). By definition, this is a web application or web site that fits on a single web page, aiming to provide a smoother user experience than a desktop application [1]. In SPA all of the necessary code (HTML, JavaScript, CSS) is invoked when loading the single page, or the needed resources are dynamically loaded and added to the page when requested, usually in response to user's interactions with the application. The SPA is not reloaded at any time in the process, nor does it refer to another page. For the developing of the SPA is used AngularJS.

AngularJS is a powerful JavaScript framework for developing dynamic web applications [2]. A big plus for AngularJS is that it has a built-in set of ready-to-use modules that make it easy to develop a SPA. The routing functionality that Angular provides helps to create SPA when multiple views can be loaded into single page, based only on a URL parameter.

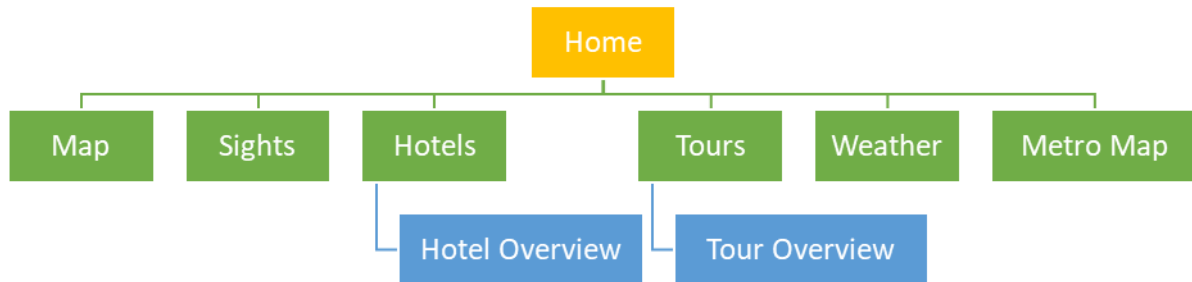
Other important part and conception in the developing of a SPA are the views – in this case is used **ngView** module of Angular, which defines what contains each page or part of the app and how it behaves when the user interacts with it [2].

The **ngView** directive specifies where in the HTML code of each page will be placed in its structure [2]. This directive adds the template of the current routing in the main structural file. In simple words – it takes the file that is mentioned for the current routing (location) and injects it into the structure of the ngView directive.

When the HTML is ready, routes must be configured. It is used the service **\$routeProvider** from **ngRoute** module. For each path it has to be determined **templateUrl** (the path to the template that loads in place of ngView when it enters the given location) and **controller** (the logic that will be performed for the particular template).

If a situation comes when the user goes to a different location, thus enter another link, we can make sure by using **otherwise** function which serves as the default route. In this case, it is set to "/".

On Fig. 1 is shown the application's structure:



**Fig. 1.** Application Structure

### 3. Program implementation

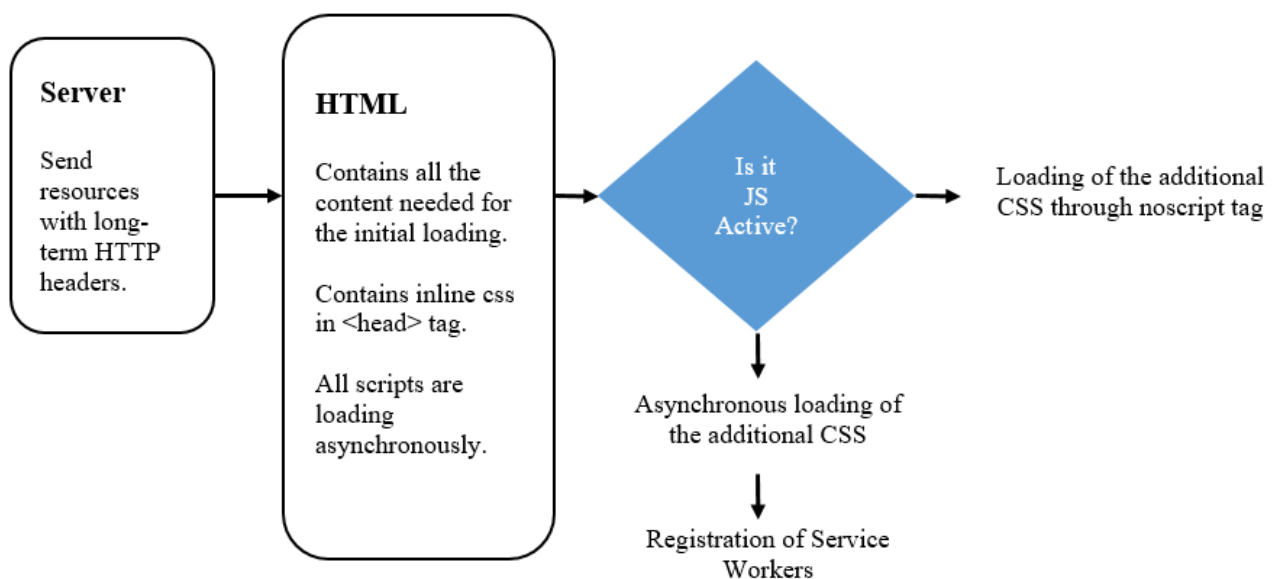
At the initial opening of the web application the primary goal is to show a meaningful content to the user as quickly as possible. Generally, the architecture needs to do a few things:

- to prioritize the initial loading by allowing the service worker to cache the application shell so that subsequent visits do not require the shell to be invoked from the network;
- to “lazy” load everything else or to load it in a background.

The application shell uses minimal HTML, CSS and JavaScript, which creates the user interface (UI). The shell should:

- Load fast;
- Be cached;
- Dynamically display content [3].

The app shell is the secret to good and reliable performance. It's supports the UI locally and dynamically retrieves the content through the API.



**Fig. 2.** Initial page load and Non-Service Worker load of the page [3]

To achieve this is necessary:

The server sends the HTML content that the client can render and uses long-term HTTP cache headers for browser without service workers support. Provides files by using hashes to allow versioning and easy updates for a later period of the cycle of the application.

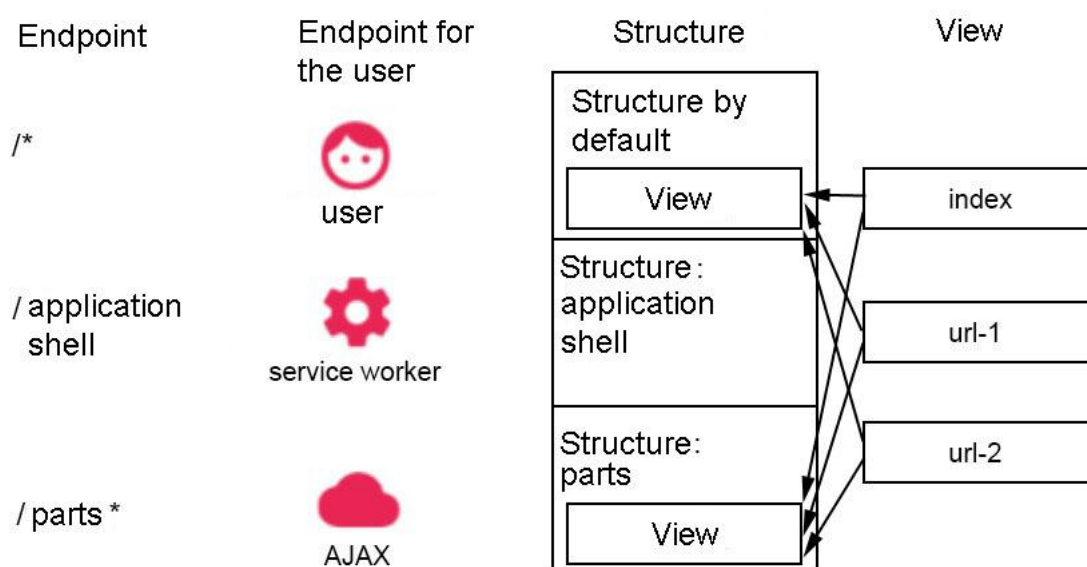
The pages contain inline CSS styles in one <style> tag in the <head> tag of the document to provide a quick display of the application's shell styles. Each page loads asynchronous JavaScript needed for the current view.

Service Worker keeps a cached version of the application shell so that on a subsequent visit, the shell itself is fully loaded by the service worker cache unless there is an update available on the on the network [3].

### Server part

The application's server component processes the content and the presentation layer separately. The content can be added to the HTML structure that would end up in static page rendering or in dynamic processing and loading.

This architecture looks like this:



**Fig. 3.** User and server endpoints [3]

The endpoints are defined in 3 parts: the user who opens particular URL (index, url-1), the app shell (service worker) and the HTML parts.

Each endpoint has a controller through which the views (which are actually some parts of the HTML code) are inserted into the final page. In the application, Angular 's data-binding functionality also plays a big role in making it easier to embed the application shell and synchronize the individual parts.

Initially, the user sees a static content page. This page registers the service worker if it is supported by the browser, that caches the application shell and everything it is dependent on (CSS, JavaScript, etc.).

The application shell itself then plays the role of SPA by using JavaScript to send requests into the content for a specific URL. Those XHR requests are sent to `/parts*` which returns small pieces of HTML, CSS and JavaScript, needed to show this content.

For the browsers which are not supporting service workers should be given the opportunity to switch to an alternative. In this case, the application will not be cached and will not be available

offline. Our research system is uploaded and can be found at the following link: <http://georgigyurov.github.io>.

#### **4. Main metrics for researching**

##### **Request and response headers size**

Most web sites already use compression when downloading text content, as it provides significant productivity improvements. Unfortunately, HTTP/1.1 does not support compression of the HTTP headers that are attached to each HTTP request and response. SPDY and following HTTP/2 are designed to address this gap by using different types of compression.

SPDY uses the DEFLATE algorithm for basic purposes where HTTP/2 uses HPACK that is specifically designed to compress headers. It uses previously defined tokens, dynamic tables and Huffman encoding [4].

It is possible to see a difference in the size of the generated headers by reviewing a non-content request.

##### **Response message size**

The server response message is consisting of response headers and encoded content from the response.

##### **Number of TCP connections for page load**

The Browsers improve performance in HTTP 1.1 by increasing the maximum number of host connections from two to six or more. This allows simultaneous downloading of single page at the cost of additional TCP connections and SSL handshakes. Increasing the concurrent connections allows network bandwidth to be used more efficiently because it reduces blocking of the requests.

SPDY and HTTP/2 support concurrency for single TCP and SSL connections by using multiplexing to allow more than one request per unit of time to send and receive data for a single connection.

##### **Page load time**

This parameter indicates when the web page is fully loaded and ready for use. In most cases, this is a reasonable measure of the page speed as well as the site visitor sees it.

NGINX server that supports SSL certificate and HTTP/2 on localhost is chosen for research environment. The main advantages that NGINX holds and due to them it is chosen as a web server and research environment are high performance and an effective basis for HTTP processing of a desktop, mobile and API traffic. NGINX has a built-in ability to handle many connections by processing thousands of them per second.

#### **5. Performing of the research and analysis of the results**

Once the server is set up with the deployed application, the data is gathered from the metrics that are marked in the previous chapter. Firefox is used as the primary browser for testing and its switching functionality between SPDY and HTTP/2, and for some of the tests is used also Google Chrome's DevTools. The Boolean value of "network.http.spdy.enable.http2" is switched between true and false, depending on which protocol is going to be used. Another alternative to switch between the two protocols is a setting in the server's configuration file. Each test was performed in a brand fresh installation of Firefox (or Chrome) with an empty browser cache. The aim is to show the main and important differences between the two protocols.

## Request and response headers size

URL заявка: https://localhost:2016/#/	Метод за заявка: GET	Отдалечен адрес: 127.0.0.1:2016	Код на състоянието: 200 OK	Версия: HTTP/1.1
Филтриране на заглавките				
Заглавки на отговора (0.235 KB)				
Accept-Ranges: "bytes"				
Connection: "keep-alive"				
Content-Length: "27978"				
Content-Type: "text/html"				
Date: "Sat, 25 Jun 2016 13:16:29 GMT"				
Etag: ""576d49ae-6d4a""				
Last-Modified: "Fri, 24 Jun 2016 14:54:38 GMT"				
Server: "nginx/1.10.1"				
Заглавки на заявката (0.346 KB)				
Host: "localhost:2016"				
User-Agent: "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.11; rv:47.0) Gecko/20100101 Firefox/47.0"				
Accept: "text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8"				
Accept-Language: "bg,en-US;q=0.7,en;q=0.3"				
Accept-Encoding: "gzip, deflate, br"				
Connection: "keep-alive"				
Pragma: "no-cache"				
Cache-Control: "no-cache"				

Header size for SPDY

URL заявка: https://localhost:2016/#/	Метод за заявка: GET	Отдалечен адрес: 127.0.0.1:2016	Код на състоянието: 304 Not Modified	Версия: HTTP/2.0
Филтриране на заглавките				
Заглавки на отговора (0.173 KB)				
Date: "Sat, 25 Jun 2016 13:11:48 GMT"				
Etag: ""576d49ae-6d4a""				
Last-Modified: "Fri, 24 Jun 2016 14:54:38 GMT"				
Server: "nginx/1.10.1"				
X-Firefox-Spdy: "h2"				
Заглавки на заявката (0.409 KB)				
Host: "localhost:2016"				
User-Agent: "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.11; rv:47.0) Gecko/20100101 Firefox/47.0"				
Accept: "text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8"				
Accept-Language: "bg,en-US;q=0.7,en;q=0.3"				
Accept-Encoding: "gzip, deflate, br"				
Connection: "keep-alive"				
If-Modified-Since: "Fri, 24 Jun 2016 14:54:38 GMT"				
If-None-Match: ""576d49ae-6d4a""				
Cache-Control: "max-age=0"				

Header size for HTTP/2

Fig. 4. Header sizes for SPDY & HTTP/2

It is possible to see the difference in the generated sizes for the request and response headers when using the two protocols. The two pictures show the size of the request and response headers where the HTTP/2 response is considerable smaller than SPDY, while the HTTP/2 request is bigger. HTTP/2 should also have a smaller request header size due to its use of the HPACK algorithm.

## Response message size

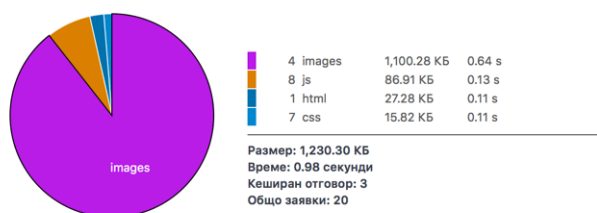


Fig. 5. Response message cached – HTTP/2 & SPDY

The body of the returned response from the server is composed of the header of the response and the encoded content of the response. Given the fact that HTTP/2 should create the smallest headers shouldn't it always have the smallest response message.

In the obtained results the size of the image files is the same for both protocols, and the difference is with the other files where HTTP/2 has a predominance in the size and the time. The difference in text resources will not always be in favor of HTTP/2, since optional padding bytes are added which can be attached to the HTTP/2 DATA frame. Padding is used to determinate the exact frame size and is provided to mitigate specific HTTP attacks. For example, attacks where the



compressed content includes hacked text and hidden data [5]. Padding isn't used for image resources because they already have a compressed format that does not contain attacker controlled plain text [6].

### Number of TCP connections for page load

TCP connect is the time it takes to create a TCP connection to the web server or the proxy. If a secure connection (SSL) is used, then its time is also included. When the browser accesses a web page, it must create one or more TCP connections. This could be a long-term operation even with normal and unsecure HTTP. Using of Keep-Alive reduces this problem by reusing TCP connections for multiple HTTP requests.

It is measured the parameter **connectStart** which is taken when an initial connection to the server is made. If TLS or SSL are used, then **secureConnectionStart** will start when the handshake starts securing the connection. **connectEnd** shows when the connection to the server is completed.

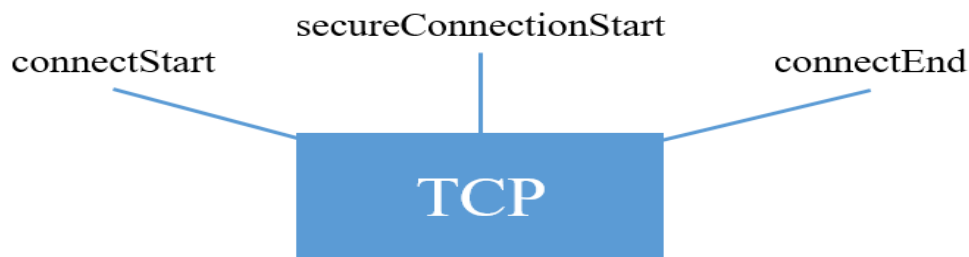


Fig. 6. TCP timeline

Table 1. TCP connectStart for SPDY & HTTP/2

	connectStart	connectEnd	Duration	responseStart	responseEnd
<b>SPDY</b>	<b>197.96</b>	197.96	71.95	260.96	269.91
<b>HTTP/2</b>	<b>157.42</b>	157.42	31.34	183.10	188.76

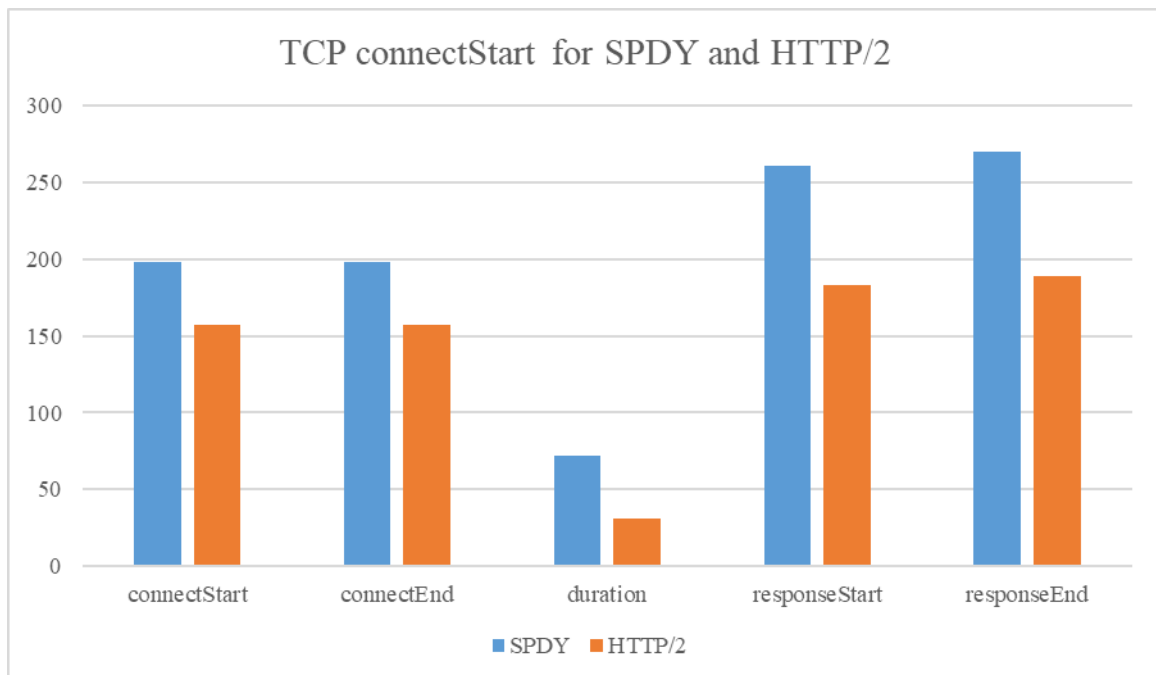
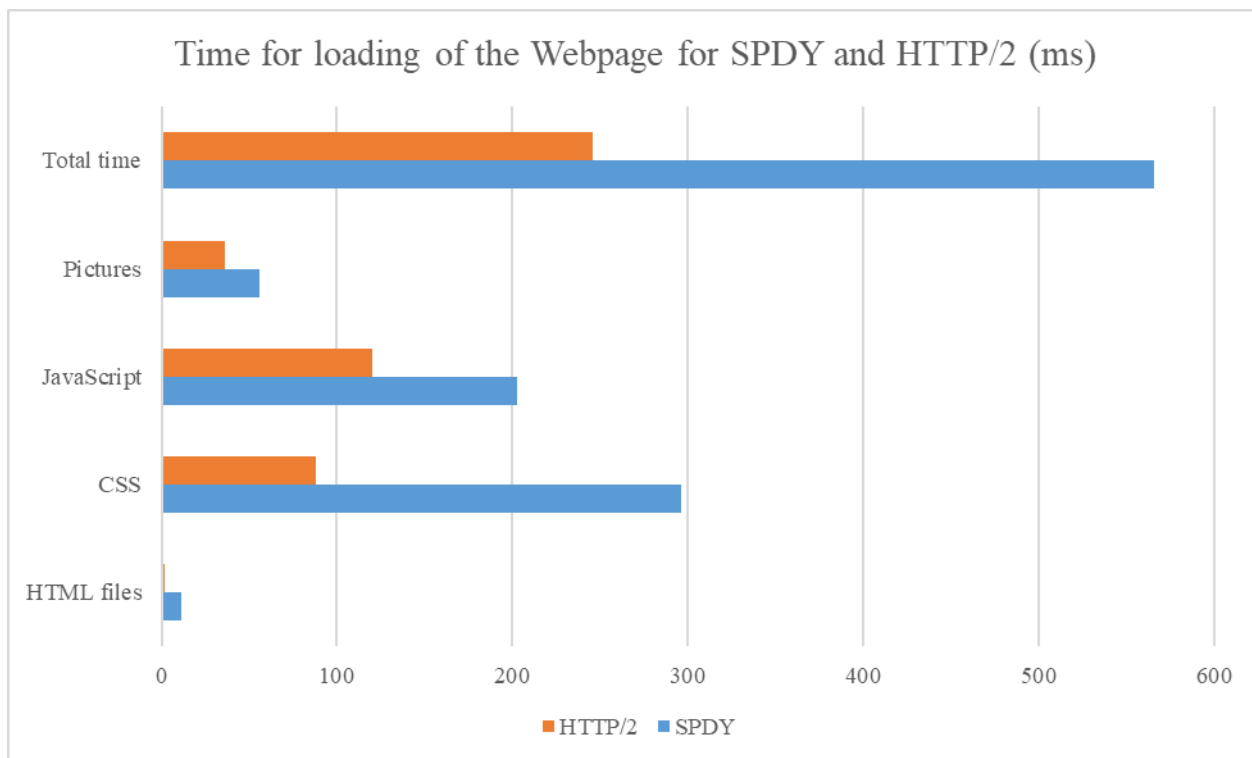


Fig. 7. TCP connectStart for SPDY & HTTP/2

From the parameters **connectStart** and **duration** it can be seen that the more connections are made, the longer is the time for their execution to the server increases, i.e. the time and the number of connections for SPDY and HTTP/2 is significantly different where HTTP/2 has an advantage.

### Page load time

It is measured the time for the full page loading and ability to work with. In most cases this is a good speed measure, because this speed is also the perception of the user which uses the page.



**Fig. 8.** Page load time for SPDY & HTTP/2

Taking into account the data from column Time and Timeline tabulated, it is easy to conclude that HTTP/2 has a tremendous advantage over SPDY, and even the proposed system is not subjected to performance optimizations.

Such optimizations would be concatenation of script and css files, minification of these files, css sprites, etc. In some cases, such optimizations would also have a counter-effect because of the size that will be the ultimate result of their merging.

## 6. Conclusion

Multiplexing support added to SPDY and HTTP/2 reduces the number of connection that need to be set to download a web page. As an additional benefit, the web servers will not need to maintain so many active TCP connections when HTTP/2 support becomes more widespread and used.

The page load time for SPDY is worse due to the fact that the protocols has no headers compression and additional TCP connections and SSL handshakes that are a must. For more sophisticated pages, the SPDY and HTTP/2 speed advantage should be even better.

Based on the research, we can conclude that HTTP/2 is significantly faster than SPDY, despite the size of its headers. For most cases the connection to the Internet is asymmetric – the upload speed is less than the download speed. This means that all of the saved data for uploading have more impact than the equivalent savings for the downloaded data.

HTTP/2 is likely to bring significant performance improvements compared to pure HTTPS and SPDY. However, the use of a padding in response messages is a possible area of concern where there can be an exchange of performance and security.

Most users are not concerned about what protocol is used by particular web site or application – they just want it to load quickly and work as expected. The benefits of switching to HTTP/2 depend mainly on the structure of the particular web site and what percentage of its users are using modern browsers.

## References

- [1]. Single-Page Applications, <https://msdn.microsoft.com/en-us/magazine/dn463786.aspx>, Nov. 2013
- [2]. Guide to AngularJS Documentation, <https://docs.angularjs.org/guide/introduction>
- [3]. Osmani, A., Gaunt, M., Instant Loading Web Apps with an Application Shell Architecture, <https://developers.google.com/web/updates/2015/11/app-shell>, Nov. 2015
- [4]. Al-sharif, I., HTTP 1.1, SPDY 3.1 and HTTP2 Protocols Performance Comparison., Jul. 2015
- [5]. Belshe, M., Peon, R., Hypertext Transfer Protocol Version 2 (HTTP/2), <http://http2.github.io/http2-spec/#padding>, May 2015
- [6]. A simple Performance Comparison of HTTPS, SPDY and HTTP/2, <https://blog.httpwatch.com/2015/01/16/a-simple-performance-comparison-of-https-spdy-and-http2/>, Jan. 2015

### For contacts:

Assist. Prof. Vladimir D. Dimitrov, PhD  
Computer Systems Department  
Technical University of Sofia  
E-mail: vldimitrov85@gmail.com

Eng. Georgi Gyurov  
Computer Systems Department  
Technical University of Sofia  
E-mail: georgigyurov7@gmail.com



# СИСТЕМА ЗА ПРОСЛЕДЯВАНЕ НА АВТОМОБИЛ ПОСРЕДСТВОМ МОБИЛНИ ТЕХНОЛОГИИ

Георги И. Георгиев, Милко Т. Маринов

**Резюме:** GPS проследяването на автомобили става все по-актуална тема в наши дни. Нерелационните бази от данни набират все по-голяма популярност сред потребителите и разработчиците. Реализирането и намирането на методи за увеличаване на производителността на система с използването на нерелационна база от данни в независима мобилна система за проследяване на автомобил е целта на този проект. Разработена е цялостна система, чрез която може да се проследява автомобил в реални условия. Модулното устройство записва и предава сензорни данни, текущо местоположение и други данни по безжичен път до отдалечен сървър чрез мобилни технологии.

**Ключови думи:** Следене на автомобил, нерелационни бази от данни, MongoDB, GPS, GPRS.

## Vehicle tracking system using mobile technologies

Georgi I. Georgiev, Milko T. Marinov

**Abstract:** GPS tracking of cars is becoming an increasingly hot topic these days. NoSQL databases are gaining in popularity among users and developers. The realization and finding methods to increase system performance with the use of NoSQL of databases in an independent mobile tracking system is the goal of this project. A system has been developed to track the car location in real conditions. The modular device records and transmits sensor data, current location, and others wirelessly to a remote server through mobile technologies.

**Keywords:** Vehicle tracking, NoSQL, MongoDB, GPS, GPRS.

### 1. Увод

Използването на съвременни методи за запис на хетерогенни данни с висока честота на предаване е водеща тема сред IoT общността [1]. Ограниченията пред релационните бази от данни все повече подтиква разработчиците да избират нерелационните решения поради тяхната широка гама на приложимост.

Целта на настоящата статия е да се представят някои от съвременните методи за проследяване на автомобил, отчитането на неговото местоположение, скорост и други сензорни данни. За реализирането на цялостната система са използвани модерни подходи [2], [3], технически средства и услуги [4], [5], [6]. Проучването показва нови методи за запис на данни на отдалечен сървър, работа с различен набор от сензори и модули за локализация, обработка на данни, изграждане на двустранни канали за пренос на данни посредством мобилна мрежа (GPRS [7]) и работа с хетерогенни данни с висока честота на предаване.

Изграждането на система, справяща се с разнороден тип данни, е с висок приоритет при решението на проблема. За нуждите на изследването бе избрана нерелационна база от данни MongoDB [5], която е водеща сред нерелационните бази от данни поради редица нейни преимущества. MongoDB е поредната по ред система за управление на база от данни (СУБД), която изследваме за работа с хетерогенни данни с висока честота на предаване. По своята производителност тя превъзхожда многократно релационните системи. Основните характеристики на MongoDB, които бяха водещи при избора на подходяща СУБД за целите на проекта, са следните: динамични схеми, структурирани данни в JSON формат, индексирание на полета, съвместимост с JavaScript (употреба на MapReduce операции), ограничени колекции (кръгова опашка).

В таблица 1 е сравнено времето, необходимо за запис на данни в нерелационна и релационна СУБД. Тъй като времето за запис е от голяма важност за предаването на хетерогенни данни с висока интензивност, бе необходима система, която да поддържа високи скорости на запис. MongoDB поддържа високи скорости на запис и четене, което я прави предпочитан избор за изграждането на цялостната система.

**Таблица 1.** Време за запис при различни СУБД

	Време за запис в секунди				
Брой записи	100	1000	5000	25000	50000
MongoDB	0,1	0,2	0,2	0,5	1,2
MS SQL	0,1	1	10	100	160

За целите на изследването бе създаден модул, който да покрива поставените критерии за скорост на пренос на данни, честота и обработка на различни по вид данни. Изискванията към системата подробно са описани в секция 3. Устройството бе изградено чрез микрокомпютърна система с набор от сензори и модули, обработващи и предаващи данните по GPRS мрежа.

## 2. Съществуващи решения

В днешно време съществуват многобройни системи със значителен брой функции и възможности, но, поради характера си на изработване, поддържане и реализирани възможности, са трудно приложими за целите на нашия проект.

В [8] е представена система, която намира частични решения на поставените проблеми, но не изпълва напълно изискванията (описани в секция 3). Използването на Wi-Fi модул решава техните проблеми, но не и нашите, тъй като е необходимо изпращане на данни от всяка точка на света. Изградената система намира приложение само при локални условия с наличие на Wi-Fi свързаност.

В [9] авторите прилагат добри похвати за комуникиране между мобилното устройство и сървърната част, но така разработена системата има своите недостатъци, което я прави неприложима за нашите изследвания. Използването на кратки съобщения, предавани по мрежата (SMS), може да доведе до конфликтни ситуации, тъй като съобщенията, предавани по мрежата, е трудно да се трасират дали са пристигнали и понякога биват загубени.

В [10] се демонстрира подобен подход за изпращане на данни по мрежата посредством кратки съобщения (SMS). Разликата е в използването на готов сървър, което намалява възможността за настройване и програмиране. Използването на скъпи модули за комуникация увеличава стойността на представения проект, докато при разработването на нашата система не на последно място бе взета под внимание и себестойността на целия проект.

Работата, показана в [11], демонстрира подобна реализация на система за предаване на данни по безжичен път. Разликата в GSM модулите показва как една система може да бъде изградена по различен начин. Подборът на изграждащите модула компоненти не е подходящ за мобилна система, поместена в малък корпус, както при нас.

Колегиумът в [12] е разработил система за пренос на данни в реално време, въпреки различните ситуации на пътни условия, в които може да попадне тяхното устройство. Тази функционалност може да бъде от полза в някои ситуации. Възможността за изпадане на системата в несъстоятелност е взета под внимание в реализирането на нашата система и е добавена като функционалност.

Всяка една от изброените системи има свои предимства и недостатъци, но поради характера си на приложение не покриват поставените критерии (описани в секция 3).

В таблица 2 е представен сравнителен анализ на основните параметри на подобни апаратни системи:

**Таблица 2.** Сравнителен анализ на подобни апаратни системи

Устройство	Програмируемост	Мобилност	Локализация	Честота	Батерия
<b>GPS прос. устр.</b>	Не	Да	15м	1200 MHz	1200mAh
<b>Мобилен телефон</b>	Да	Да	10м	1400 MHz	2200mAh
<b>GPS/GSM устр.</b>	Да	Да	7м	1500 MHz	7000mAh

Първото устройство представлява GPS проследяващо устройство, което няма възможност за програмиране и има малка вградена батерия, която ограничава приложението му. Второто устройство е стандартен мобилен телефон, който има достатъчно функционални възможности, но не е рентабилен да се препрограмира и използва за такава цел. Освен това максималният му период на работа на батерията е не по-голям от 20 часа. Последното устройство е създадено специално за да покрива поставените условия, да е лесно програмируемо и да има дълъг период на работа.

### 3. Изисквания към системата

Предаването на хетерогенни данни към отдалечен сървър по мобилната мрежа е основната задача на цялостния проект. За да е възможно това, системата трябва да отговаря на следните изисквания:

- Предаване на хетерогенни данни с висока честота на предаване;
- Отдалечено предаване на информация посредством мобилна мрежа (GPRS);
- GPS локализиране;
- Предаване на сензорни данни (температура, атмосферно налягане);
- Запис на данни в JSON формат;
- Цикъл за проверка на актуалността на данните;
- Обработка на данни на отдалечен сървър (CRUD функционалност);
- Запис на данни в нерелационни бази от данни (MongoDB);
- Независимо хранване на модулното устройство (работа на батерия за период от поне 4 дни);
- Малки габарити на крайното устройство (корпус на модул за заключване от автомобил).

На база на поставените изисквания бе изграден модул, който поддържа следната функционалност:

- GPS позициониране;
- GPS проследяване при превишена скорост;
- Циклична проверка за актуални данни;
- Рестартиране на апаратното осигуряване в случай на несъстоятелност;
- Предаване на сензорни данни в JSON формат.

Освен стандартното GPS позициониране (координати – географска дължина и ширина), системата записва скоростта, надморската височина, брой сателити и температура.

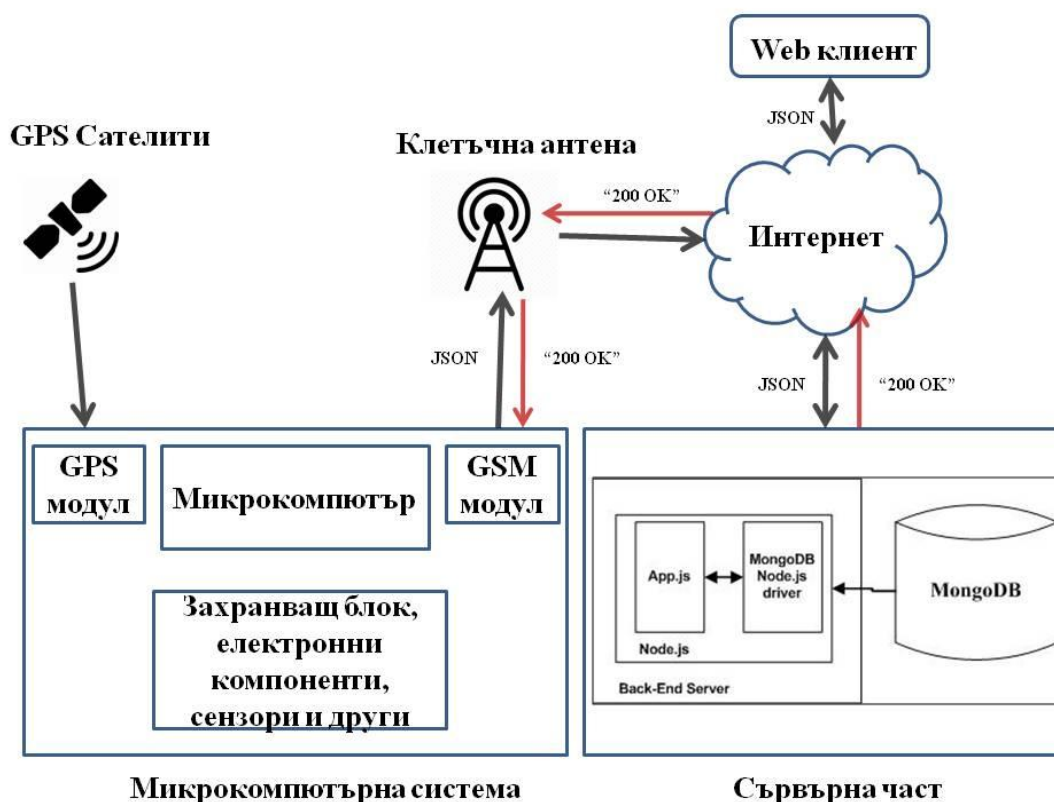
Допълнителна функционалност, изразяваща се в предупреждение за превишена скорост, е добавена като пример за увеличаване на функционалните възможности чрез промяна на софтуера, инсталиран на микрокомпютъра. Лесната замяна на кода, управляващ

вътрешните процеси на микрокомпютъра, увеличава приложението на системата. Поставянето на независимо захранване на мобилното устройство бе водещо изискване при разработването на системата.

В системата има изработена схема със съобщения за обратна връзка. Докато модулът не получи информационно съобщение, той не изпраща следващия пакет с данни. Реализирането на функционалността е изобразено на фигура 1.

#### 4. Реализация

Изградената система се състои от програмно и апаратно осигуряване. Програмното осигуряване служи за обработка и запис на постъпилите данни от мобилното устройство, поставено в автомобила. Апаратната реализация представлява микрокомпютърна система, работеща с набор от сензори и модули за трансфер на данни посредством GPRS мрежа. Общата архитектура на системата е представена на фигура 1.



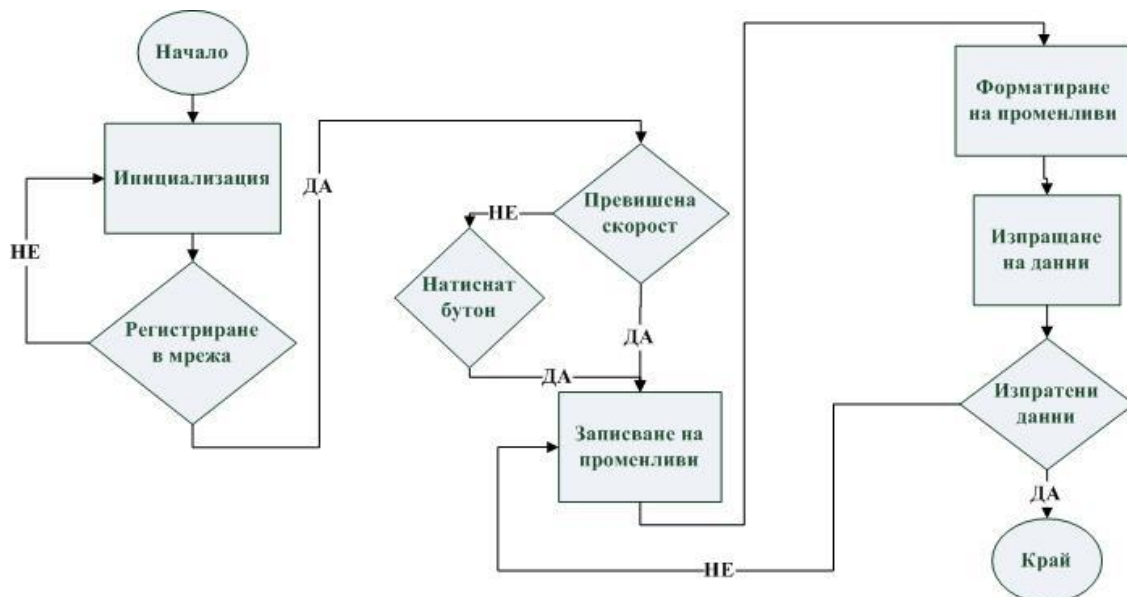
Фиг. 1. Обща архитектура на системата.

Софтуерната част на проекта е изградена от два основни компонента:

- Сървърна част с помощта на Node.js и Express. Сървърната част обработва постъпилите данни от хардуерното устройство в JSON формат и записва информацията в MongoDB. Програмният код е написан на Java програмен език;
- Софтуер, обработващ вътрешните процеси на микрокомпютъра, командите за инициализация на модула, функциите, обработващи и изпращащи данни по GPRS мрежата и други системни команди.

Преди да постъпят в сървъра данните биват предварително обработени в нужния формат (JSON). Хардуерното устройство предварително обработва данните с актуалните моментни стойности, след което ги преобразува в подходящ формат и изпраща по мрежата. На фигура 2 е представена обработката на вътрешните събития и изпращането на данните.





**Фиг. 2.** Обработване на събитията от мобилното устройство.

При наличие на два отделни програмни модула за обработка на данни (в мобилното устройство и на сървър) е необходимо взаимното им комуникиране. В системата са използвани многобройни системни проследяващи съобщения, уверяващи следващата команда, че предишната е приключила успешно. Пример за това е обратно постъпилото съобщение: „200 OK“ (от сървър), което мобилното устройство транслира като: „SEND OK“. Съобщението получено и обработено означава, че данните са изпратени и записани в сървър и може да се изпратят нови данни.

Методът за комуникация между устройството и сървър е чрез HTTP Post заявка. При изграждането на системата бе взет под внимание начина на комуникация, за това бе избран подходящ GSM модул, който поддържа съответния метод. След успешно свързване на модула със сървър започва предаване на данни в JSON формат. Пример за JSON заявка е представен на фигура 3:

```

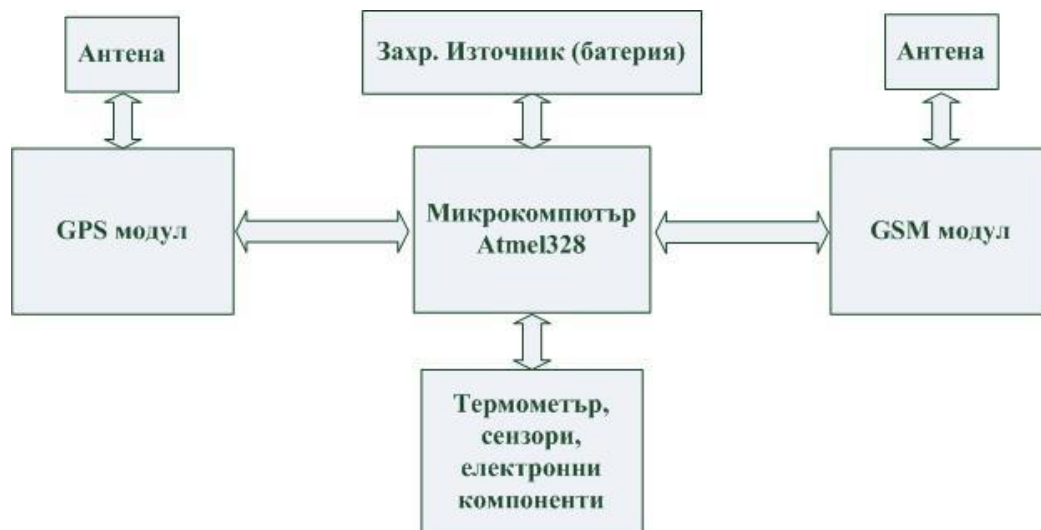
{
  "latitude": "43.854115",
  "longitude": "25.960155",
  "speed": "3.13",
  "altitude": "50.40",
  "satellites": "8",
  "temperature": "32",
  "overspeed": "no",
  "create_date": {
    "Sdate": "2018-08-24T16:11:13.177Z"
  }
}

```

**Фиг. 3.** Примерна заявка за запис

След постъпването на данните в сървър, те биват обработени и изпратени за запис в базата от данни.

На фигура 4 е представена блок схема на апаратното осигуряване.



**Фиг. 4.** Апаратно осигуряване. Блок схема

След включване на захранването, модулът стартира вътрешните проверки и инициализации, показани на фигура 2.

## 5. Резултати

След цялостното изграждане на системата бяха направени редица проверки и тестове за достоверност на данните. След включване на захранването, устройството успешно се локализира и актуализира текущите данни. При натискане на бутона за изпращане системата проверява редица параметри, описани в следна последователност:

- Проверка за реакция на устройството;
- Проверка дали е регистриран в GPRS мрежата;
- Регистриране в GPRS мрежата (ако не е предварително регистриран);
- Свързване на устройството със сайта;
- Стартиране на процедура по изпращане на данни към сървър (HTTP Post метод);
- Проверка за обратен отговор за записани данни („200 OK“);
- Затваряне на връзката към сайта;
- Очакване на нова команда за изпълнение.

Освен чрез бутон за въвеждане на данни, системата следи за превишаване на предварително зададена стойност на скоростта, при превишаване на параметъра започва цикъл на автоматично предаване на данни с интензитет 1 секунда.

При проверка на работоспособността на системата бяха отчетени времената за изпълнените на команди, описани в таблица 3.

**Таблица 3.** Времена на изпълнение на команди

№	Команда			
	Команда 1	Команда 2	Команда 3	Команда 4
<b>1</b>	2200ms	300ms	200ms	1000ms
<b>2</b>	2970ms	300ms	200ms	1000ms
<b>3</b>	3000ms	200ms	230ms	1000ms
<b>4</b>	2300ms	300ms	220ms	1000ms
<b>5</b>	2450ms	900ms	200ms	1000ms

(Команда 1 - време за изпращане на данни към сървър; команда 2 - време за свързване към сървър; команда 3 - време за изпълнение на post заявка; команда 4 - време за следваща реакция.).

След извършените проверки системата е в готовност или за ръчно стартиране на изпращането на данни, или за автоматично при превишена скорост.

Размерите на изграденото мобилно устройство бяха съобразени с началното условие (корпус от модул за централно заключване на автомобил). Габаритите на модула предразполагат към по-лесно вграждане в автомобил и съответно трудното му локализиране от крадци. С цел повишаване на часовете на работата на устройството бе необходимо добавянето на допълнителни батерии извън корпуса на устройството, което е отчетено като недостатък на модула. На фигура 5 е представен самият програмируем модул.



**Фиг. 5.** Корпус и размери на устройството.

## **6. Заключение**

След продължителна работа на системата и отчитане на данни в реални условия бе установено, че модулът се инициализира и локализира всеки път при подадено захранване. Независимото захранване успешно поддържа модула на линия за период до 5 дни. Бутонът за запис на данни отчита всяка реакция и стартира процедура по изпращане на данни към сървърната част, която завършва успешно със запис на актуална информация в базата от данни. Цикълът за предаване на данни при превишена скорост сработва всеки път при нарушено условие (по подразбиране скорост по-висока от 56 км/ч) и изпраща текущото местоположение на автомобила. Температурният сензор на устройството отчита коректно стойностите на околната среда.

Проведените изследвания дават увереност към разширяване на функционалността на системата и разработването на още два модула за предаване на данни за бъдещи изследвания.

Бъдещите разработки са насочени в инсталирането на сървърна част на едно хардуерно устройство в територията на университета, където в момента има инсталирана друга NoSQL база от данни (HyperTable [13]) с цел изграждане на однородни сървърни системи с различни бази данни. Изградените системи ще позволят бъдещи изследвания с цел сравнителен анализ и методи за подобряването на времето на записване на данни в NoSQL бази от данни от мобилно устройство.

## **Благодарности:**

Този доклад се публикува със съдействието на проект 2018-EEA-01 „Методи за събиране, организация, достъп, обработка и анализ на големи обеми от структурирани и

неструктурирани данни“, финансиран от фонд „Научни изследвания“ на Русенски университет „Ангел Кънчев“.

### Литература

- [1]. Watanabe, S., A. Nakamura. Modern Approaches for Intelligent Information and Database Systems, Vietnam, Springer, 2018, pp. 541.
- [2]. Klettke Meike, U. Storl, S. Scherzinger. Datenbanksysteme für Business, Technologie und Web (BTW 2015), Bonn, Gesellschaft für Informatik, 2015, pp. 237.
- [3]. Markey, P., G. Clynch. A performance analysis of WS-\* (SOAP) and RESTful Web Service for Implementation Service and Resource Orientated Architectures. Dublin Institute Of Technology ARROW@DIT. 2013. № 5, pp. 1-9.
- [4]. Cantelon, M., M. Harter, T. Holowaychuk, N. Rajlich. Node.js in Action, Shelter Island, Manning Publications Co, 2014, pp. 11.
- [5]. Mikowski M., J. Powell. Single Page Web Applications: JavaScript end-to-end, USA, Manning Publications Co, 2013, pp. 432.
- [6]. Maity B., S. Soumya, N. Debnath. Challenges of implementing data warehouse in MongoDB envirement. Journal of Fundamental and Applied Sciences, 2018, № 10, pp. 7.
- [7]. Yan X., J. Yang, W. Lu. Design of Remote GPRS-based Gas Data Monitoring System. IOP Conf. Series: Earth and Environmental Science, 2018, № 108, pp. 5.
- [8]. Rajatha, U., W. Sheema, R. Sheetal, B. BiAmeena, M. Tejaswini. Intelligent Parking Based on Geolocation Using IoT. 3<sup>rd</sup> National Conference on Image Processing, Computing, Communication, Networking and Data Analytics, 2018, № 8, pp. 143-146.
- [9]. Kunal, M., S. Mandeep, J. Neelu. Real Time Vehicle Tracking System using GSM and GPS Technology- An Anti-theft Tracking System. International Journal of Electronics and Computer Science Engineering, 2012, № 8, pp. 5.
- [10]. Sharanpreet, K., S. Dheerendra. Optimization in GPS-GSM Technologies used in Vehicle Tracking. International Journal of IT and Knoledge Management, 2015, № 8, pp. 39-42.
- [11]. Dr. Mohamed, Z. GSM-GPRS Arduino Shield (GS-001) with SIM 900 chip module in wireless data transmission system for data acquisition and control of power induction furnace. International Journal of Scientific & Engineering Research, 2014, № 5, pp. 5.
- [12]. Patel, P., S. Deshmukh. Geo-Location Big Data Based Data Mining Architecture Using MongoDB For Collaborative E-Initiative Based Crowd-sourced Traffic Management System. International Journal of Advanced Research in Computer Science, 2013, № 4, p148-152.
- [13]. Marinov, M., P. Vitliemov, E. Popova. Towards Big Data and Internet of Things as Key Aspects of Energy Efficiency. TEM Journal, 2017, № 3, pp. 427-435.

#### За контакти:

маг. инж. Георги И. Георгиев  
катедра „Компютърни системи и технологии“  
Русенски университет „Ангел Кънчев“  
E-mail: ggeorgiev@uni-ruse.bg

# NVME SSD ДИСКОВЕ И ТЯХНОТО ИЗПОЛЗВАНЕ В ОБУЧЕНИЕТО НА ИТ СПЕЦИАЛИСТИ

Явор З. Дечев

**Резюме:** Задачата на настоящото изследване е да се анализира производителността на NVMe SSD дискове. В резултат на изследването е направено обосновано предложение в какви направления на обучението на ИТ специалисти да се използват компютърни конфигурации, използващи изследваните устройства.

**Ключови думи:** SSD, NVMe, M.2, HDD, SATA

## NVME SSD drives and their use in the education of IT specialists

Yavor Z. Dechev

**Abstract:** The topic of this research is to analyze the performance of NVMe SSD drives. As a result of research, a well-grounded proposal is made for the different kind of directions of education of IT specialists to use computer configurations whiths the studied device

**Keywords:** SSD, NVMe, M.2, HDD, SATA

### 1. Въведение

Към настоящия момент SSD дискове се използват все повече при асемблиране на компютърна техника. Поради високата си скорост за четене и запис те все повече изместват класическите магнитни дискове. Най-широко приложение намират като системни дискове в работни станции, основни устройства за съхранение на данни в портативни устройства и при високоскоростни RAID масиви.

За разлика от магнитните дискове технологиите за производство на SSD, архитектурата на устройствата и интерфейса за връзка към тях и алгоритмите за работа са в постоянен процес на развитие.

### 2. Изложение

Задачата на настоящото изследване е да се анализира производителността на отделните типове компютърни архитектури, използвани в учебните кабинети при използването на софтуер за десктоп виртуализация.

SSD (Solid State Drive) са енергонезависими носители на информация. Те представляват масив от флаш памет, управляван от флаш контролер. Най-често се използват NAND. Клетките с памет могат да бъдат на няколко нива: SLC(Single level cell) – запомня се 1 бит в клетка; TLC (MLC-3) – запомня 3 бита в клетка или MLC (Multi-level cell). Към момента се произвеждат 72-слойни флаш памет за SSD. 3D флаш паметите осигуряват висока плътност – по-голям капацитет запомняни данни в по-малко физическо пространство. Те притежават по-висока скорост за четене и запис и занижено енергопотребление. 3D технологията намалява смущенията между отделните клетки на паметта, а от там и намалява и възможността за компрометиране на записаните данни. Основен техен недостатък са по-високите производствени разходи и по-сложен контролер за управление [4,5].

Скоростта за обмен на данни между SSD и дънната платка се ограничава от физическия интерфейс на устройството: SATA, eSATA, **PCI Express**, Fibre Channel, SAS, type-C или ATA/IDE.

Към настоящия момент се използват следните интерфейси:

- **SATA** - интерфейсът е универсален за включване както на HDD, така и на SSD устройства. Към момента е актуален SATA revision 3.2 с поддържаща скорост на записване 16 Gbit/s.
- **PCI Express** - използва се за включване на устройства директно към дънната платка. Интерфейсът за дискове може да поддържа до 16 линии (x16). **PCI Express Version 4.0** поддържа скорост за предаване на данни 16.0 GT/s за всяка линия. Разлика в поддържаната скорост може да се получи от различните слотове, в които се поставя дискът. Интерфейсът се използва и за включване на хардуерни RAID контролери със SSD масиви.
- **M.2 (Next Generation Form Factor (NGFF))**. Интерфейсът намира най-широко приложение в портативните устройства. Конекторът позволява директно включване на SSD дънната платка. Основно предимство пред стандартното PCIe свързване е малкият форм-фактор на слота и на включеното към него устройство. Съществуват различни стандарти M.2, в зависимост от физическата големина на модула и от поддържаните протоколи. Създадени са както M.2 SATA устройства, така и M.2 PCI Express. Последните поддържат максимална производителност PCI Express x4 (4 GB/s) [3].
- **Fibre Channel** – оптичният интерфейс се използва за вътрешна връзка между сървъри и за изграждане на мрежи за съхранение на данни, използващи SSD.
- ATA/IDE, mSATA, eSATA намират ограничено приложение при остарели компютърни конфигурации.

С цел повишаване скоростта на запис при SSD е добавена командата TRIM. Чрез нея при изтриване на файл операционната система няма да маркира пространството като неизползваемо, а ще го изтрие. Така при презапис върху един сектор няма първоначално да се изтрие предишната информация, а след това да се запише новата. Недостатък на командата е невъзможността за възстановяване на данни от SSD.

Протоколите за достъп до SSD са AHCI и NVMe. AHCI (Advanced Host Controller Interface) е контролер, създаден за обслужване на магнитни дискове. Той е оптимизиран за работа с ротационни дискове с висока латентност, а не за флаш памети с ниска латентност. AHCI поддържа 1 опашка с до 32 команди.

NVM Express (NVMe - Non-Volatile Memory Express) представлява контролер за PCI Express SSD. Проектиран е за работа с енергонезависима скалируема флаш памет. Използва опростен набор от команди (до 13). Предимството на архитектурата на NVMe е липсата на регистър за съхранение на командата за четене, с което се намалява допълнително латентността на диска. Допълнително контролерът поддържа до 64K (65536) броя опашки с 64K (65536) команди във всяка. Множеството опашки гарантират пълното натоварване централния процесор при обработка на данните. При портативните устройства намалената латентност допълнително води до намаляване времето за работа на диска и следователно намаляване на енергопотреблението и увеличаване на времето на живот на батерията [2,3].

Платформите AMD – Ryzen и чипсетите на Intel X97/X99 поддържат няколко PCI Express линии директно към централния процесор. Това позволява, използвайки NVMe, SSD устройството да бъде директно свързано към CPU, изолирайки се от PCH (Platform Controller Hub) контролера на чипсета [1].

Ограничението при устройствата, работещи с NVMe протокола, е, че той се поддържа само от операционни системи Windows 8.1 или 10, Windows Server 2012, Linux с ядро 3.13, Solaris S.12, VMware release 12.2013 и по-новите от тях. Хардуерните платформи трябва да

притежават чипсет на Intel – X97/X99 или по-нов, или платформа на AMD – Ryzen или чипсет 990FX. NVMe не се поддържа от BIOS, а само от UEFI. Множеството ограничения налагат, при асемблиране на работната станция, ограничен брой комбинации от по-скъпи компютърни компоненти от висок клас, което допълнително повишава цената на компютърната конфигурация.

Предимствата на SSD пред магнитните дискове, които предполагат използването им в портативни устройства или във високо производителни компютърни системи, са:

- Висока скорост за четене и запис;
- Липса на движещи механични части;
- Ниско енергопотребление;
- Малък форм-фактор.

Същевременно недостатъците на SSD са:

- Висока цена на модула и на единица GB запомняна информация;
- Кратък живот на устройството - в зависимост от използваната технология от няколко хиляди до няколко стотин хиляди презаписвания във флаш клетката;
- По-трудно възстановяване на данни;
- Ограничена хардуерна и софтуерна поддръжка.

Посочените недостатъци на SSD определят и продължаващото използване на магнитните дискове за съхранение на информация в ниско бюджетни компютърни системи или в такива, при които бързодействието на запомнящото устройство няма водещо значение.

### 3. Изследване

Параметрите на компютърната конфигурация, използвана за изследванията, са: лаптоп HP Pavilion Power 15; **CPU** - Intel Core i7-7700HQ @ 2.80GHz Kaby Lake; Cores – 4; **L1 Instruction Cache** - 4 x 32 KB; **L1 Data Cache** - 4 x 32 KB, **L2 Cache Size** - 4 x 256 KB, **L3 Cache** - 6 MB; **RAM** - 8GB DDR4 SDRAM PC4-19200; **HDD** – 1TB, SATA; **SSD** - 256 GB SAMSUNG 960 Pro, NVMe M.2; **Video Card** - NVIDIA GeForce GTX 1050, memory - 4 GB.

Изследваният SSD е SAMSUNG 960 M.2 с големина от 256 GB. Дискът е флаш базиран и е изработен чрез технологията 3D NAND. Той поддържа NVMe Express, функциите NCQ, TRIM, Auto Garbage Collection Algorithm и 256-bit AES криптиране. Изследваният HDD е произведен от Seagate. Голямината му е 1 TB. Използва SATA III интерфейс и LBA адресация. Логическите дялове за форматиране чрез MBR Partition Scheme.

И на двата системни дяла е инсталиран Windows 10 Enterprise Edition build 17134 (64-bit). По време на тестовете единствено е стартирана вградената в операционната система защита Windows Defender.

Производителността на отделните компютърни компоненти е тествана чрез Performance Test 9.0 (64bit) (Trial) и Crystal Disk Mark 6.0.1 (64bit). Изследванията са извършени по време на конвертиране на видеофайл, инсталиране на десктоп дистрибуция на Linux в среда за десктоп виртуализация и при проверка за наличие на malware.

Първата задача е изпълнена чрез Wondershare Video Converter Ultimate Version 8.7.0 (trial). Конвертиран е видеофайл (Format: MKV; Size: 33,9 GB; Resolution: 1920\*1280). Параметрите на изходният файл са – Format: AVI; Resolution: 1920\*1280). И двата файла се намират върху изследвания диск.

Виртуалната машина е създадена в среда на софтуер за десктоп виртуализация Oracle VM VirtualBox 5.2.18. Нейните параметри са: 4 процесорни ядра, 5 GB споделена оперативна памет, 128 MB споделена видеопамет, 25 GB виртуален диск, създаден върху SSD или HDD, включен мрежов контролер. Тестването на производителността на компонентите се извършва по време на процес на инсталиране на операционна система в среда за виртуализация. Избрана е тази задача, тъй като периодът на инсталиране е по-дълъг от



времето за провеждане на отделните тестове. Допълнително натоварването на системата е увеличено с едновременната поддръжка на две операционни системи и обработването на сравнително голям единичен файл. Избрана е десктоп дистрибуция на Linux – openSuse 15.0, която използва по-голям iso файл в сравнение с повечето масови дистрибуции (3,8 GB).

Проверката за наличие на malware е извършена чрез портативния инструмент Kaspersky Virus Removal Tool 15.0.19.0. Проверен е единствено системният диск. Времето за извършване на проверката не е отчитано, поради различния обем от информация на отделните дискове.

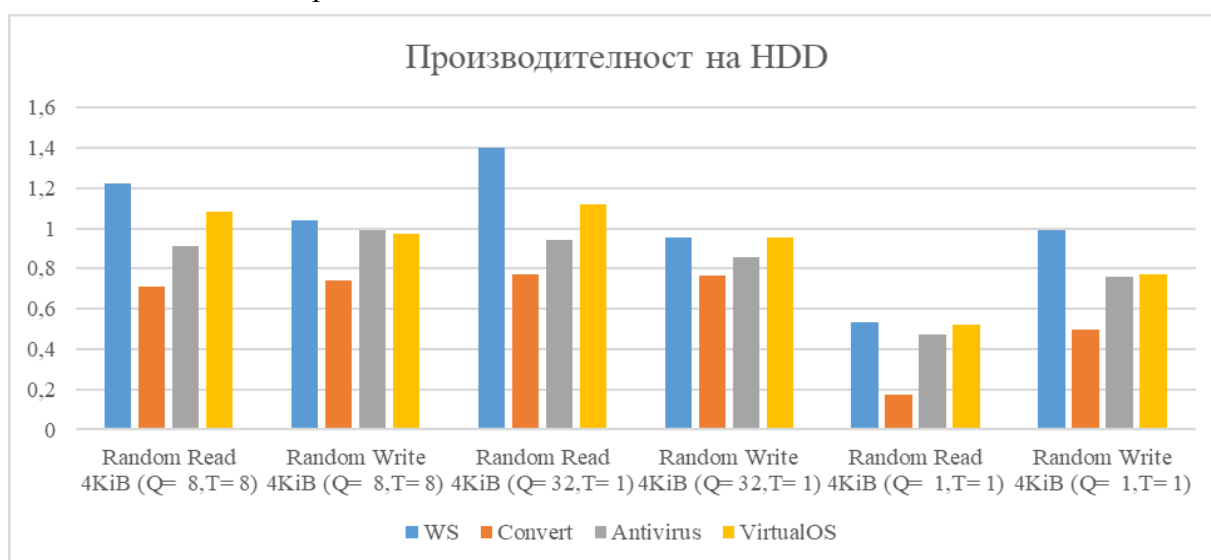
Изследванията са извършени на два етапа. Първо е проверена производителността на отделния компютърен компонент без натоварване (WS-Workstation), а след това и по време на работата с посочените софтуерни продукти.

Резултатът от проведените тестове са посочени в таблици 1, 2, 3 и 4.

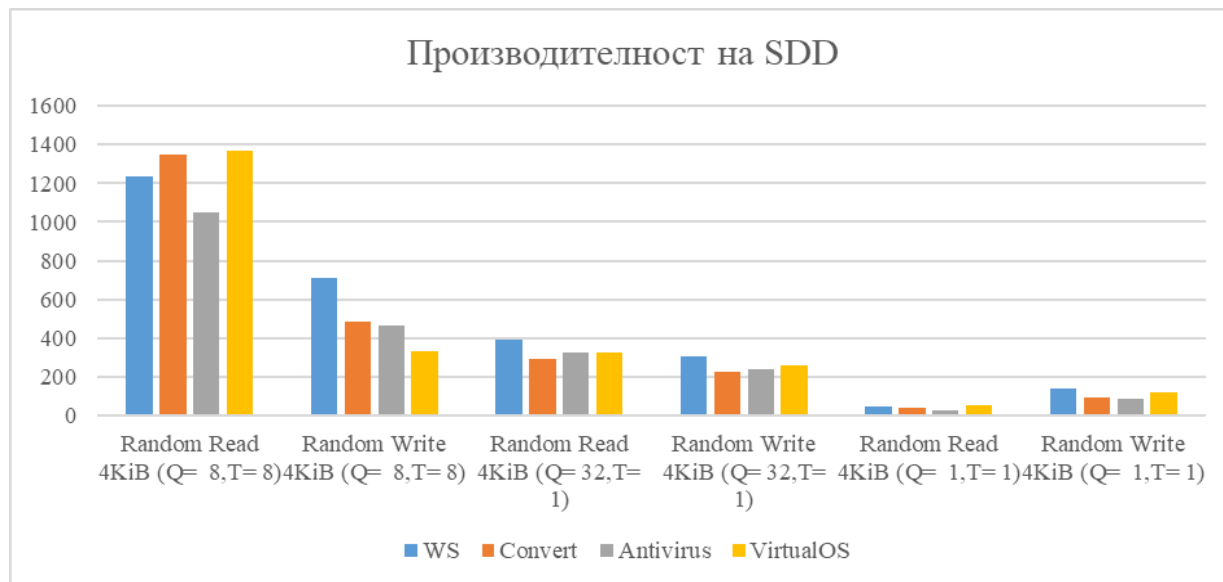
**Таблица 1.** – Резултати от тестване на производителността на дисковете с Crystal Disk Mark 6.0.1

	HDD				SSD			
	WS	Convert	Antivirus	VirtualOS	WS	Convert	Antivirus	VirtualOS
Sequential Read (Q= 32,T= 1)	179,416	86,681	145,757	160,461	3337,166	2797,48	2930,84	2812,837
Sequential Write (Q= 32,T= 1)	182,559	103,253	171,736	154,72	1190,89	1203,093	1014,976	33,637
Random Read 4KiB (Q= 8,T= 8)	1,22	0,712	0,909	1,081	1237,305	1346,695	1051,834	1368,639
Random Write 4KiB (Q= 8,T= 8)	1	0,744	0,991	0,973	710,911	485,539	465,778	328,882
Random Read 4KiB (Q= 32,T=1)	1	0,774	0,945	1,118	393,314	292,721	322,663	324,243
Random Write 4KiB (Q= 32,T=1)	0,955	0,764	0,856	0,954	304,807	225,061	237,09	260,307
Random Read 4KiB (Q= 1,T= 1)	0,534	0,174	0,474	0,521	45,925	41,717	24,168	53,141
Random Write 4KiB (Q= 1,T= 1)	0,994	0,495	0,76	0,771	136,238	92,567	85,297	119,793

Обобщените данни за производителността на дисковете, извършена с Crystal Disk Mark 6.0.1 са показани на диаграми 1 и 2.



**Диаграма 1** – Производителност на HDD



**Диаграма 2 – Производителност на SSD**

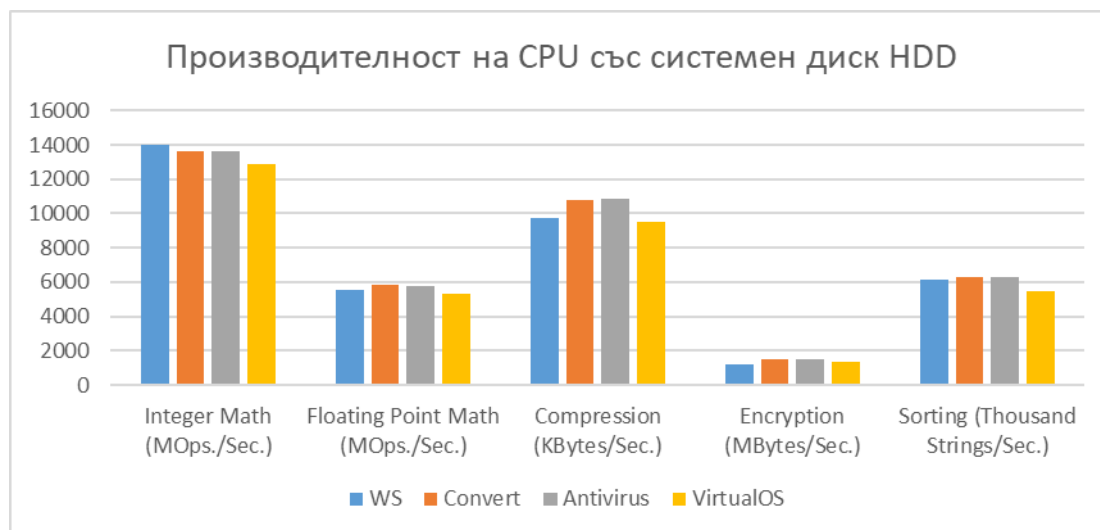
**Таблица 2. – Резултати от тестване на производителността на дисковете с Performance Test 9.0**

	HDD				SSD			
	WS	Convert	Antivirus	Virtual OS	WS	Convert	Antivirus	Virtual OS
Disk Sequential Read (MBytes/Sec.)	95,40	87,40	110,30	136,30	2468,50	1719,40	1817,00	1728,50
Disk Sequential Write (MBytes/Sec.)	52,20	51,60	29,10	75,30	1137,70	1170,40	769,80	687,30
Disk Random Seek + RW (MBytes/Sec.)	5,30	4,30	3,40	5,70	592,20	727,80	658,30	281,20

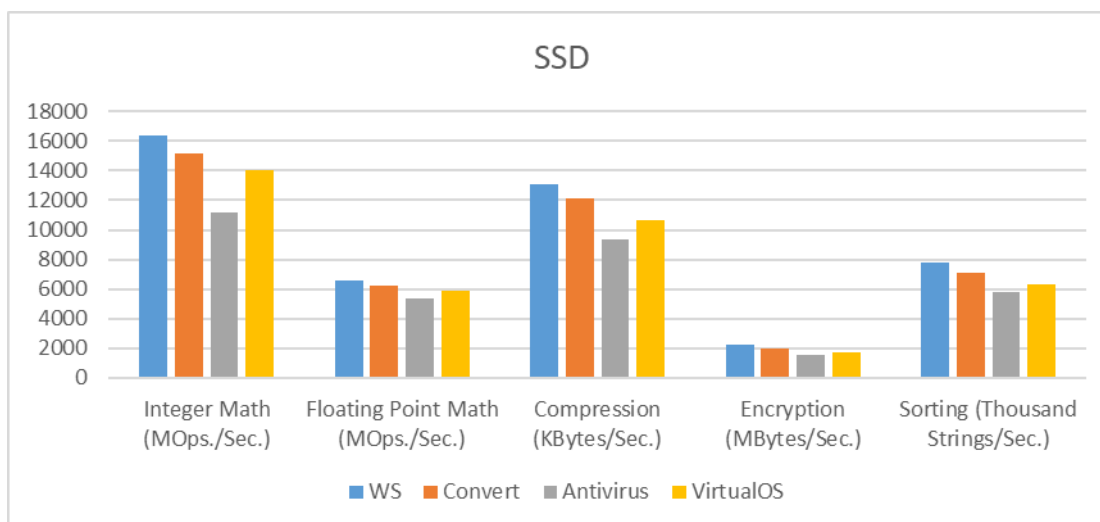
**Таблица 3. – Резултати от тестване на производителността на централния процесор с Performance Test 9.0**

	HDD				SSD			
	WS	Convert	Antivirus	VirtualOS	WS	Convert	Antivirus	VirtualOS
Integer Math (MOps./Sec.)	14034	13662	13615	12865	16414	15126	11163	14023
Floating Point Math (MOps./Sec.)	5565	5837	5750	5349	6604	6195	5322	5865
Compression (KBytes/Sec.)	9719	10748	10849	9508	13049	12159	9317	10677
Encryption (MBytes/Sec.)	1234	1493	1508	1363	2201	2010	1584	1712
Sorting (Thousand Strings/Sec.)	6180	6267	6291	5442	7789	7072	5810	6292

Обобщените данни за производителността на централния процесор, извършена с Performance Test 9.0, са показани на диаграми 3 и 4.



**Диаграма 3** – Производителност на CPU със системен диск HDD

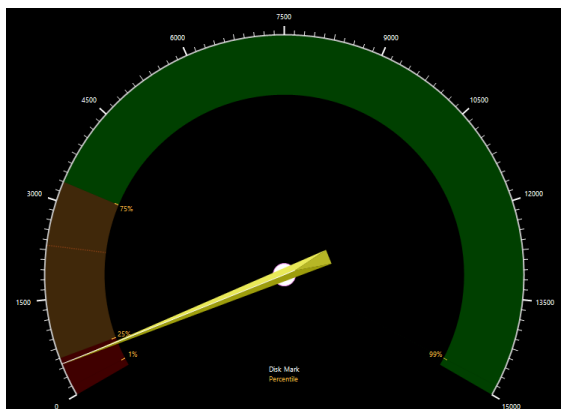


**Диаграма 4** – Производителност на CPU със системен диск SSD

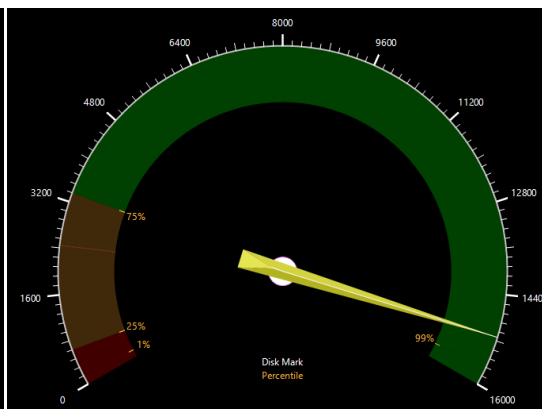
**Таблица 4.** – резултати от тестване на производителността на оперативната памет с Performance Test 9.0

	HDD				SSD			
	WS	Convert	Antivirus	Virtual OS	WS	Convert	Antivirus	Virtual OS
Memory Read Cached (MBytes/Sec.)	20546	20827	23412	21248	24566	18181	13609	20233
Memory Read Uncached (MBytes/Sec.)	8793	8941	9643	8529	14666	8734	5800	9882
Memory Write (MBytes/Sec.)	4431	4548	4877	4326	7033	4751	3783	4784
Available RAM (Megabytes)	5370	4922	4526	3070	5959	5588	5142	2891
Memory Latency (ns (lower is better))	33,50	30,10	30,50	34,00	29,50	32,00	41,80	27,50
Memory Threaded (MBytes/Sec.)	11284	11183	11027	10418	16360	16173	14502	14666

Визуалните резултати за производителността на твърдите дискове, извършена с Performance Test 9.0 са показани на фигури 2 (HDD) и фигура 2 (SSD).



Фигура 1



Фигура 2

## Резултати

- При тестване с Crystal Disk Mark 6.0.1 и последователно записване и четене разликата в скоростите варира от 6 до 32 пъти по-висока при SSD. При произволно четене и запис тя нараства от 50 до 1890 пъти по-висока при SSD, в зависимост от модела на теста.
- При работата на различните приложения производителността на SSD намалява от 10 до 25 % от номиналната стойност. При HDD производителността намалява в реда: виртуална машина - антивирусна програма - конвертиране на видеофайл. При някои от тестовите тя достига до 30 % от номиналната стойност.
- При тестване на диска с Performance Test 9.0 се наблюдават подобни резултати с тези, извършени с Crystal Disk Mark 6.0 – при последователен запис и четене от 9 до 26, а при произволен – от 49 до 194 пъти по-висока скорост при SSD, в сравнение с HDD. Отново се запазва редът за намаление на производителността на HDD при работа на различни приложения.
- При тестване на централния процесор с Performance Test 9.0, при използване на системен SSD, неговата производителност нараства от 15 до 44 % в зависимост от типа на теста.
- При работата на различните приложения производителността на CPU със системен SSD намалява в реда: конвертиране на видеофайл (с max 9 %)– виртуална машина (с max 22 %)– антивирусна програма (с max 32 %).
- При работата с приложения производителността на CPU със системен HDD, в зависимост от приложения тест, е различна. Наблюдава се както намаляване (до 12 %), така и увеличаване (до 22%).
- При тестване на оперативната памет с Performance Test 9.0 се наблюдават подобни резултати с тези извършени върху CPU. При използване на системен SSD, нейната производителност нараства от 10 до 40 %, в зависимост от типа на теста, при използване на системен HDD;
- При работата на различните приложения производителността на оперативната памет със системен SSD, намалява в реда: конвертиране на видеофайл – виртуална машина– антивирусна програма.
- При работата с приложения производителността на оперативната памет със системен HDD, в зависимост от приложения тест, е различна. Наблюдава се както намаляване (до 43 %), така и увеличаване (до 14%).

Анализирайки получените резултати след приключване на експеримента, могат да бъдат направени следните констатации:

- В зависимост от използвания тест, производителността на разглеждания SSD е от 6 до 1890 пъти по-висока от HDD.
- При липса на дефрагментиране на HDD и разполагане на файловете, които се обработват, на дял различен от системния, се увеличава времето за търсене на физическия сегмент от магнитната плоча на HDD, в следствие на което се намалява и производителността на цялата система;
- При използване на SSD се увеличава натоварването на централния процесор и оперативната памет поради големия брой обработвани опашки и по-големия обем от данни, които те обработват за единица време;
- Към 1.09.2018 г. цените на изследваните запомнящи устройства в amazon.com са: HDD – около 60 \$ (0,06 \$/1 GB); SSD – 339 \$ (1,32 \$ / 1 GB). [6]

#### 4. Изводи

- Използването на SSD дискове с поддръжка на NVMe за учебни цели е подходящо при работни станции, осигуряващи практическите занятия, свързани работа с големи по обем файлове или масиви от данни – администриране на бази от данни, създаване и редактиране на филми, музика, 3D анимация.
- Използването на единични бройки лаптопи с SSD дискове с поддръжка на NVMe е целесъобразно при обучението на докторанти и магистри, при които темата на дисертацията е свързана с горепосочените приложения.
- При закупуване на компютърна техника в държавните учреждения и учебни заведения най-често един от основните критерии е цена. Следователно високата единична цена на SSD дискове с поддръжка на NVMe и на компютърните компоненти, осигуряващи тяхната работа, определя малкия брой работни станции, които ще бъдат използвани за учебни цели.

#### Литература

- [1]. Bates, S. Donard: NVM Express for Peer-2-Peer between SSDs and other PCIe Devices. // Storage Developer Conference Santa Clara. 2015
- [2]. Ellefson J. NVM Express: Unlock Your Solid State Drives Potential. Flash Memory Summit. 2013
- [3]. M.2 SSDs: Aligned for Speed. Samsung Electronics America, Inc. 2015. <<https://www.samsung.com/semiconductor/minisite/ssd/download/documents/>>
- [4]. Micheloni, R, L. Crippa, C. Zambelli, P. Olivo. Architectural and Integration Options for 3D NAND. //Computers 2017, 6, 27
- [5]. Samsung V-NAND technology. Samsung Electronics Co., Ltd. 2014
- [6]. Amazon.com

**За контакти:**

д-р Явор З. Дечев  
катедра „Информационни технологии“  
ВВМУ „Н.Й.Вапцаров“ гр.Варна  
E-mail: dechco@naval-acad.bg

# ИЗПОЛЗВАНЕ НА ДЕСКТОП ВИРТУАЛИЗАЦИЯ ЗА ВИРТУАЛНИ ЛАБОРАТОРИИ

Делян Г. Генков

**Резюме:** Настоящата разработка има за цел да представи възможности за използване на десктоп виртуализацията за провеждане на лабораторни упражнения по компютърно ориентирани дисциплини, както и предимствата и недостатъците на този подход. Описана е една от възможните реализации, заедно с първоначалните резултати и изводи от експерименталното прилагане на този модел.

**Ключови думи:** Виртуализация, обучение, VMware ESXi, виртуални лаборатории.

## Using Desktop Virtualization for Virtual Laboratories

Delyan G. Genkov

**Abstract:** This paper aims to present the possibilities for using desktop virtualization for the laboratory exercises in computer-oriented topics, the advantages and drawbacks of this approach. One possible implementation is described, along with preliminary results and conclusions for experimental usage of this model.

**Keywords:** Virtualization, Education, VMware ESXi, Virtual Laboratories

### 1. Увод

При провеждането на лабораторни упражнения по различни дисциплини от учебния план на специалност „Компютърни системи и технологии“ в Технически университет – Габрово традиционно се разчита на физически десктоп компютри. Обикновено в една лаборатория се провеждат упражнения по няколко различни дисциплини от различни преподаватели. През годините бяха идентифицирани някои недостатъци на този подход:

- Инсталирането на софтуер, необходим за една дисциплина, може да повреди инсталиран софтуер, необходим за друга дисциплина, което обикновено се установява при опит за провеждане на упражнение по втората дисциплина и затруднява или прави невъзможно провеждането на занятието;
- Когато студент приключи изпълнението на дадено упражнение, неговите резултати остават записани локално на компютърната система. При следващо провеждане на същото упражнение с друга група, новият студент често намира резултатите наготово. Това налага преподавателят след всяко упражнение да търси и премахва получените файлове с резултати;
- Компютърните системи трябва да са достатъчно мощни, за да могат да стартират всяка необходима програма за всяка дисциплина. Това налага периодичното актуализиране на системите;
- По някои различни дисциплини се налага да се използват различни операционни системи (напр. Windows и Linux). Инсталирането на две операционни системи на един компютър понякога води до усложнения, например автоматичните актуализации на едната система могат да попречат на възможността за стартиране на другата и до невъзможност за провеждане на занятието;
- Понякога студентите са повече от компютърните системи, което налага част от тях да не участват активно в провеждането на упражнението;

- Случайната или нарочна повреда на дадена операционна система налага преинсталирането на цялата конфигурация, заедно с приложния софтуер, което отнема време.

Част от проблемите бяха частично заобиколени чрез промяната на начина на провеждане на упражненията, като се използваха предварително подготвени виртуални машини, стартирани на локалния компютър чрез безплатния софтуерен продукт – Oracle VirtualBox. Това донякъде намали проблемите, свързани с повреждането на файлове и операционни системи, както и локалното запазване на получени резултати, но въведе някои нови проблеми:

- Едновременното копиране на предварително подготвен файл с виртуална машина на 10 работни места от един файлов сървър може да отнеме значително време от необходимото за провеждане на упражнението, дори и след промяната на мрежовата свързаност на 1 Gbit/s. Проблемът може да се заобиколи, ако файловете се копират предварително, но това изисква доста време и усилия от преподавателя преди занятието;
- Едновременното стартиране на няколко виртуални машини на локалния компютър значително намалява производителността му;
- При софтуера Oracle VirtualBox разделянето на интерфейсите в различни мрежови сегменти става чрез задаване на различно име на мрежата в настройките на адаптера. Това не е видимо от студента и често води до объркване и липса на комуникация между виртуалните машини.

## 2. Изложение

При направеното проучване за съществуващи решения на проблема бяха идентифицирани доста подходи, голямата част от които представляваха специализирани софтуерни платформи, изискващи разработката на специфични софтуерни симулационни модули за всяко упражнение и дисциплина.

Един възможен подход, използващ KVM за виртуализация и собствен Web базиран интерфейс за конфигурация, е представен в [1]. Той е ориентиран главно към мрежови лаборатории, реализирани с устройства на Cisco Systems. Въпреки очевидните му предимства, този подход не е подходящ за нашата постановка, тъй като не беше идентифицирана възможност за осигуряване на лабораторните упражнения по всички необходими дисциплини.

Представената в [2] виртуална лаборатория по дисциплината "Организация на компютъра" представлява софтуерна симулация на процесите в компютърната система и е специализирана по конкретната дисциплина.

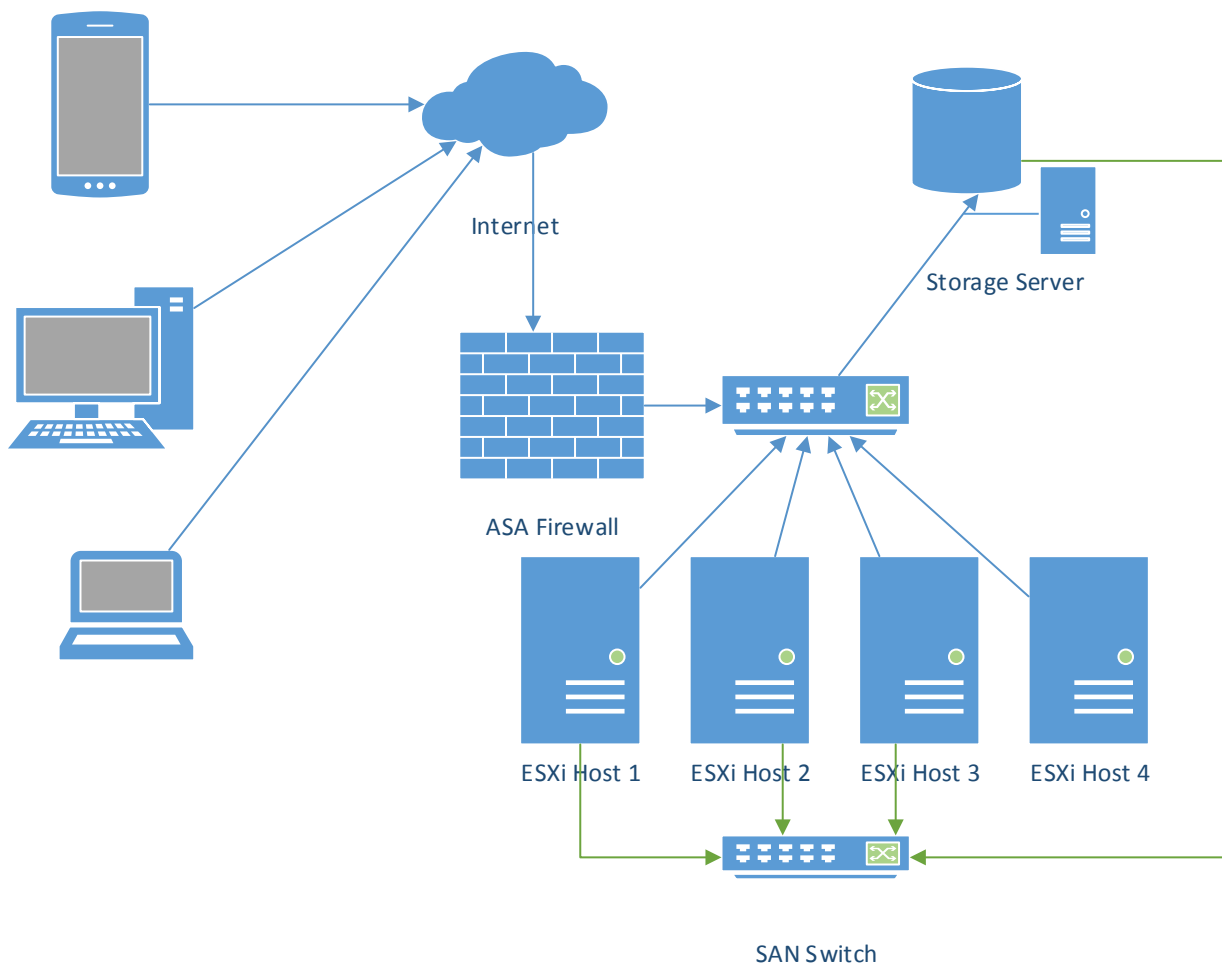
Най-близо до нашите цели се оказа представената в [3] система за автоматизация на виртуална лаборатория при обучението по информационни технологии, която е публикувана през 2011 година, като концепциите и продуктите, използвани в нея, към момента са доста остарели и недостъпни.

В рамките на изпълнение на проект № BG16RFOP002-2.009-0043 „Развитие на капацитета и интернационализацията на продуктите на Клъстер Айтос /ITOS/“ с финансовата подкрепа на ЕС чрез Европейския фонд за регионално развитие, Технически университет – Габрово, като член на клъстера, участва в спецификацията и разработката на развойна лаборатория за облачни и виртуализационни технологии. За целите на проекта беше доставена и инсталирана в град Габрово, извън територията на Университета, инфраструктура, състояща се от следните компоненти:

- 4 броя сървъри HP DL380Gen 9, 2 X 8 Core Intel Xeon E5-2620v4, 256 GB RAM, 2 X 1,2 TB 6G SAS HDD, 32 GB Flash Media Kit, 2 X 800 W Hot Swap PSU;

- 1 бр дисков масив за съхранение на данни HP DL380Gen 9, 2 X 8 Core Intel Xeon E5-2620v4, 256 GB RAM, 12 X 4 TB 6G SAS HDD, 2 X 800 W Hot Swap PSU;
- Комутатор на мрежата за данни HP Aruba 3810M 16 SFP Switch;
- Комутатор на мрежовата свързаност Cisco Catalyst 2960-X 48 GigE;
- Защитна стена Cisco ASA 5508-X with FirePOWER Services;
- VMware VSphere Standard Server + VCenter;
- VMware Horizon View + Client license pack.

На фигура 1 е показана структурата на платформата.



**Фиг. 1.** Структура на платформата

Четири физически сървъра и системата за съхранение на данни имат по четири физически мрежови контролера за трансфер на данни със скорост по 1 Gbit/s, обединени в общ логически канал (Etherchannel) с обща скорост 1 Gbit/s. Те, заедно с контролера за отдалечен достъп до конзолата на сървъра (iLo), са свързани към мрежовия комутатор Cisco Catalyst 2960-X, като са разделени в различни VLAN мрежи. Всеки сървър има и по два 10 Gbit/s мрежови контролера, свързани към комутатора за трансфер на данни HP Aruba 3810M, като общата скорост към дисковия масив на всеки сървър е 20 Gbit/s. Връзката към Интернет се осъществява през защитна стена Cisco ASA 5508-X, която регулира достъпа на отдалечените потребители през Интернет и криптира данните чрез SSL VPN. Отдалечените потребители трябва да имат инсталиран софтуера за VPN достъп Cisco AnyConnect VPN, чрез

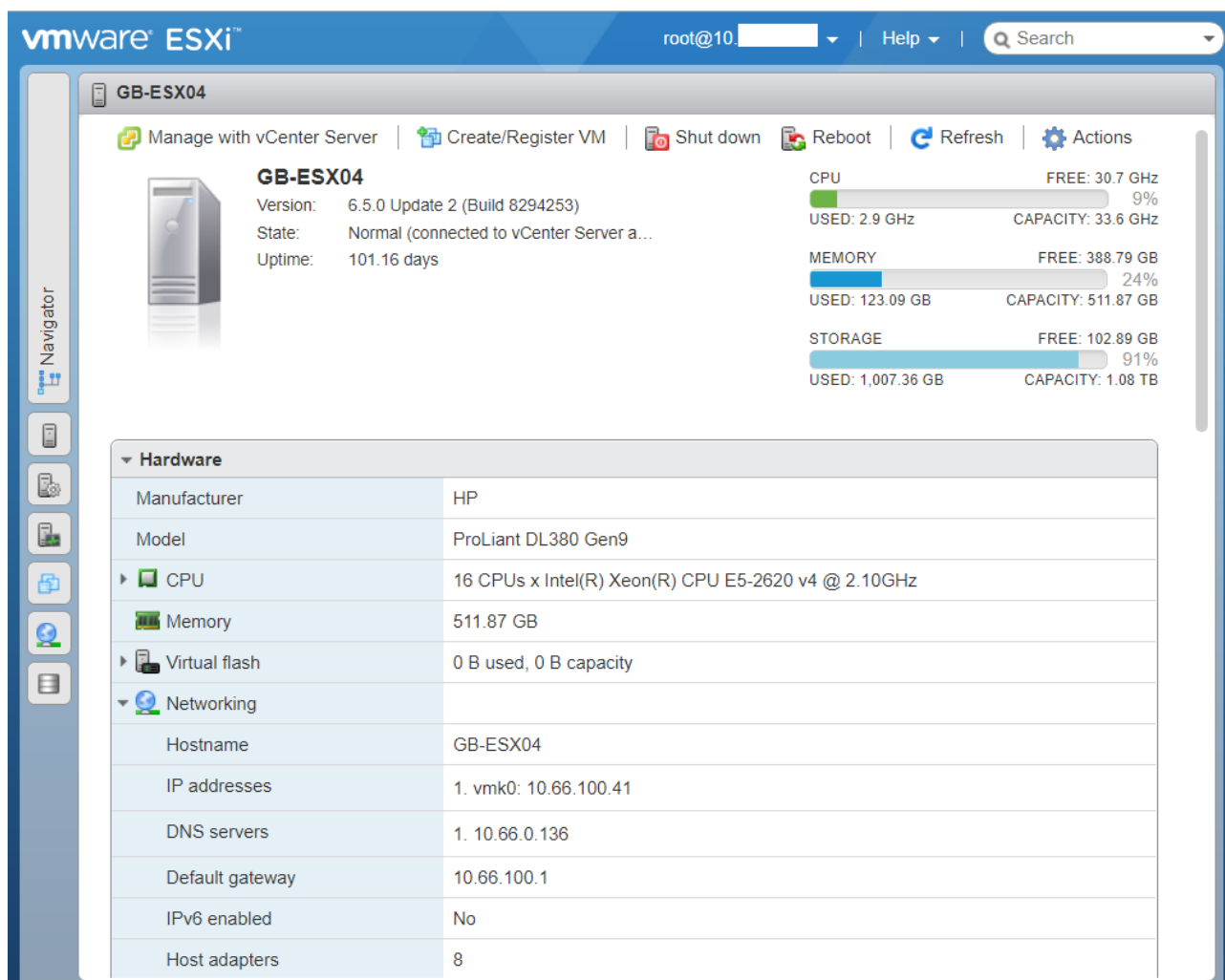


който осъществяват автентикацията и достъпа до платформата. Клиентът поддържа платформи Windows, Linux, MacOS, iOS и Android.

На всеки от четирите сървъра е инсталиран хипервайзор VMware ESX 6.5.0. Координацията и наблюдението на ресурсите се осъществява чрез VMware vCenter, инсталиран на две виртуални машини върху операционна система Microsoft Windows 2012 Server. Потребителите на системата за виртуална десктоп инфраструктура, техните права, достъп до ресурси и други параметри се управляват чрез VMware Horizon. Повече сведения за платформата могат да бъдат намерени в [4].

С помощта на платформата беше реализирана виртуална десктоп инфраструктура, позволяваща изпълнението на виртуални машини през Web browser на произволен клиент.

На фигура 2 е показан екранът на един ESX сървър със стартирани 17 виртуални машини – 5 броя Windows Server 2016, 5 броя Windows 7, 5 броя Windows 10 Professional и 2 броя Linux Ubuntu 64 bit.



Фиг. 2. Натовареност на сървър

Вижда се, че натовареността на процесора на сървъра е 9%, а на паметта му – 24%, което означава, че в тази конфигурация на един сървър ще могат да се стартират поне 60 машини едновременно, а на цялата платформа – над 200.

Към момента на написване на този текст все още е в процес на тестване и оценяване натоварването на процесори и памет на устройствата при реално провеждане на лабораторни

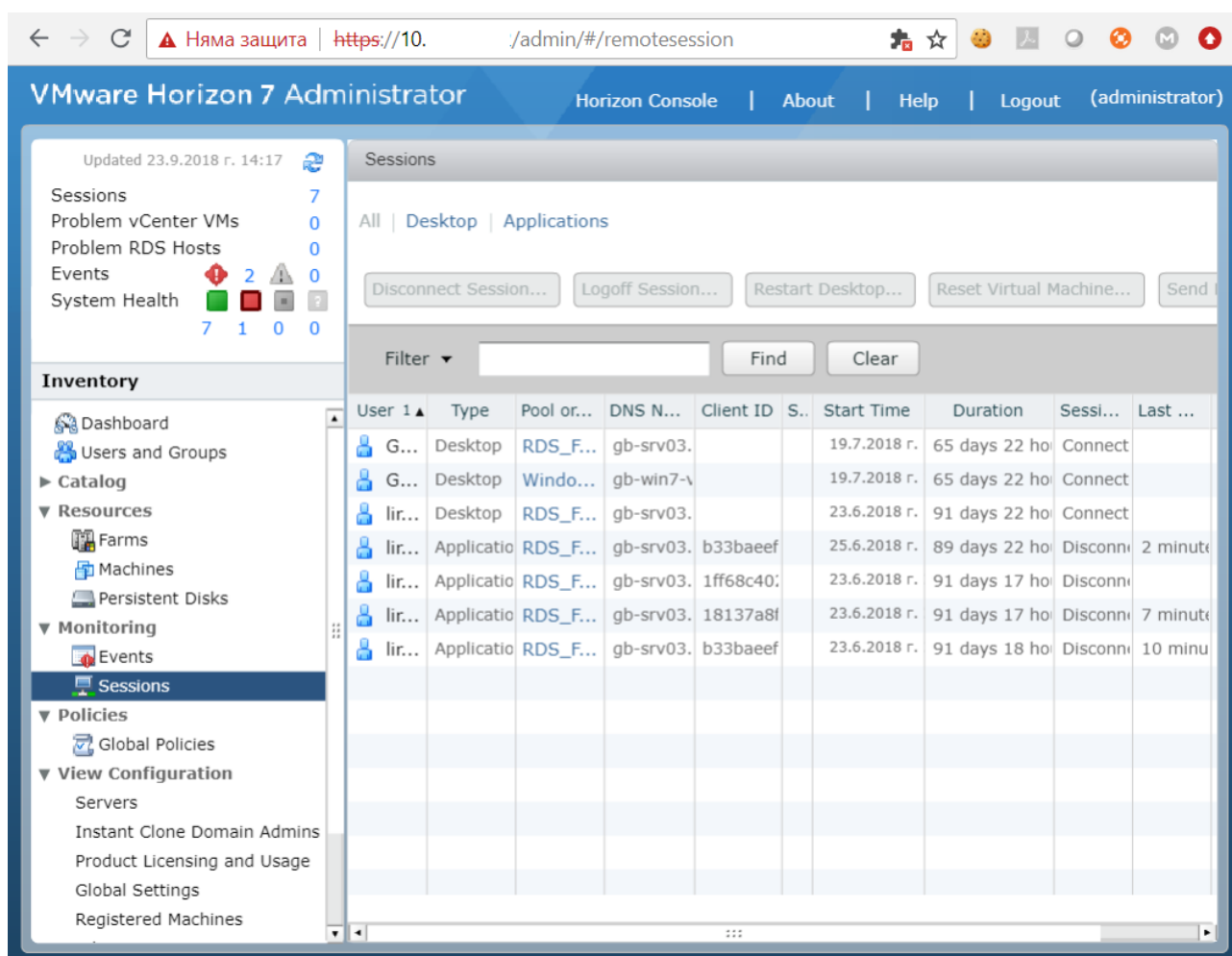
упражнения с по-тежки виртуални машини, както и необходимата скорост на мрежовия достъп до системата.

### 3. Резултати

След конфигурацията на необходимите параметри във VMware Horizon – потребители, права, пароли за достъп, виртуални машини и други, стартирането на виртуалните машини беше тествано на различни клиенти, между които:

- десктоп и лаптоп компютри с Windows 10 и Ubuntu Linux операционни системи;
- умни телефони с операционни системи Android и MacOS;
- тънки клиенти HP t520;
- HP Chromebook 11 G3;
- таблети Lenovo Yoga, Apple Airpad Air 2 и други.

От тестовите беше установено, че всяка платформа може да стартира и работи успешно виртуална машина под управлението на произволна поддържана операционна система. На фигура 3 е показан екран от платформата за виртуална десктоп инфраструктура – VMware Horizon с показани успешно изградени сесии и стартирани машини.

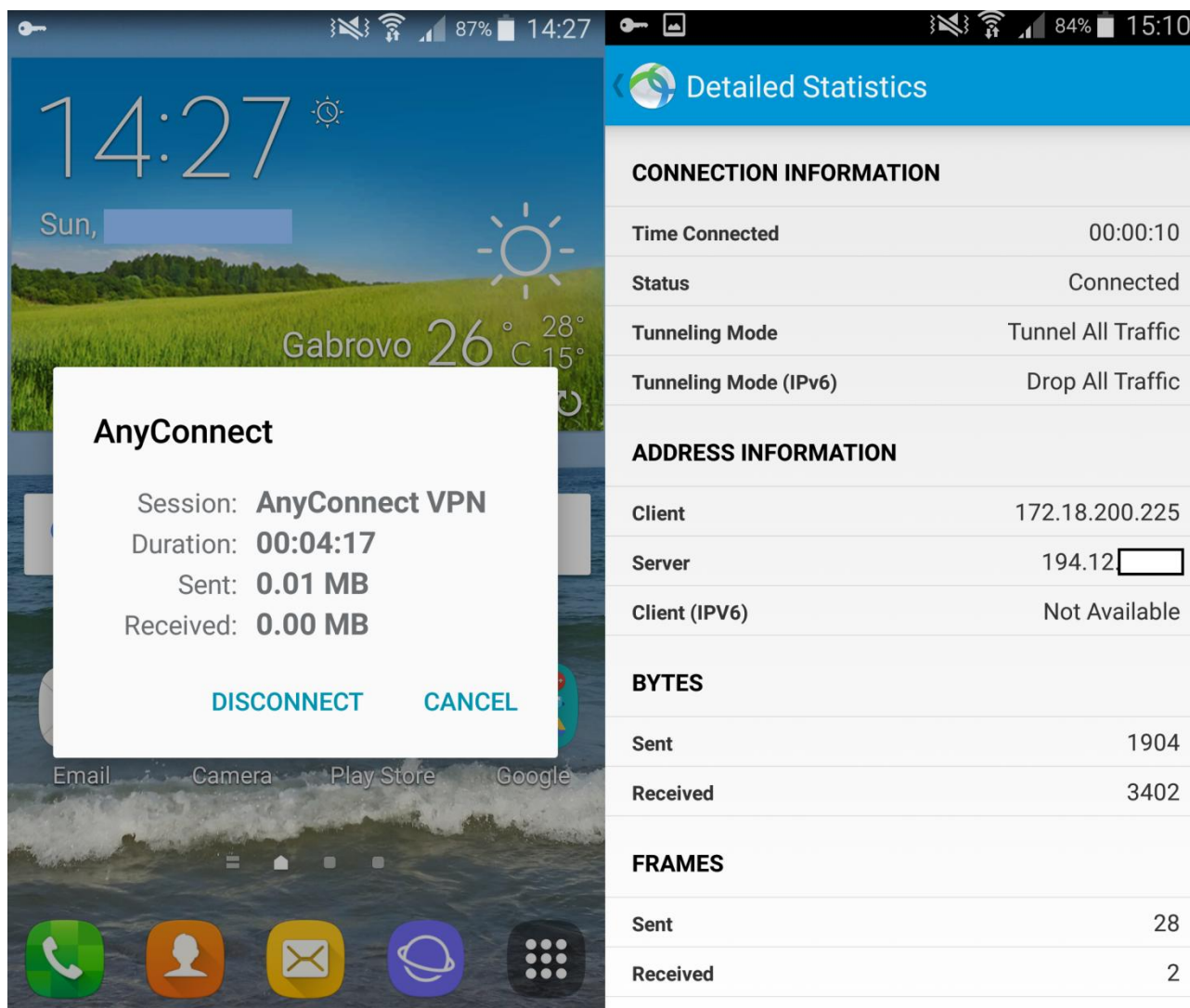


Фиг. 3. Изградени сесии към платформата

Поддръжката на различни мобилни платформи дава възможност студентите да участват в лабораторното упражнение, използвайки своите собствени терминали – лаптопи, таблети и мобилни телефони. По този начин изградената платформа дава възможност и за изпълнение

на задачите, поставени в лабораторното упражнение и извън университетските лаборатории, което може да се използва в бъдеще при осигуряване на дистанционно обучение на студенти.

На фигура 4 е показана установена VPN сесия чрез Cisco AnyConnect VPN Client от мобилен телефон под операционна система Android.



Фиг. 4. Свързване с Anyconnect VPN към платформата

#### 4. Заключение и бъдещи планове

В настоящия документ е описан един подход за изграждане на система за десктоп виртуализация, която се планира да се използва при провеждането на лабораторни упражнения на студентите от специалност „Компютърни системи и технологии“ в Технически университет – Габрово. Към момента на написване на документа системата е изградена и тествана като базови възможности за отдалечено свързване и стартиране на виртуална машина от широк набор клиентски устройства. Предварителните тестове показват, че системата е оразмерена с достатъчен капацитет, за да поеме натоварването от необходимия брой клиенти за реално провеждане на упражненията.

Предстои изграждането на виртуалните машини за провеждане на упражнения по дисциплините „Компютърни мрежи“ и „Бази данни“, тестването и въвеждането в реален учебен процес на упражненията.

Механизмът на десктоп виртуализация помага за преодоляването на редица проблеми, свързани с провеждането на лабораторните упражнения на физически компютърни системи, като инсталирането на много програми на една система, неволно или нарочно повреждане на файлове и операционни системи, както и използването на готови резултати от упражнението, получени от предишни групи. Подходът дава възможност за едновременно използване на различни операционни системи, софтуерни приложения, бързо възстановяване от откази и сринове, както и бързо и лесно запазване на резултатите за по-нататъчно използване или продължаване на дейностите по провеждане на упражнението.

Отдалеченият достъп до лабораторните упражнения дава възможност на студентите за изпълнение на задачите извън територията на университета и в произволно време от денонощието, което дава предимства при невъзможност да се посещават учебните занятия или за реализация на дистанционно обучение.

Поддръжката на много платформи и операционни системи дава възможност упражненията да бъдат стартирани и изпълнявани от мобилните устройства на студентите и по този начин намалява разходите по закупуване на стационарни устройства.

Към момента констатираните недостатъци на подхода са необходимостта от високоскоростна и гарантирана Интернет свързаност към системата за лабораторни упражнения, като има предположения, че при нестабилна или с недостатъчна скорост свързаност е възможно приложенията да работят трудно. При отпадане на свързаността към Интернет лабораторните упражнения няма да могат да се проведат. В момента се проучват възможностите за изграждане на резервирана Интернет свързаност.

### **Благодарности**

Настоящият документ е изготвен с финансовата помощ на договор № 1810Е за провеждане на научни изследвания по проект на тема: „Приложение на виртуализационни и облачни технологии в обучението“ към Технически университет – Габрово.

Апаратурата и софтуерните лицензи, използвани в документа са доставени при изпълнение на проект № BG16RFOP002-2.009-0043 „Развитие на капацитета и интернационализацията на продуктите на Клъстер Айтос /ITOS/“ с финансовата подкрепа на ЕС чрез Европейския фонд за регионално развитие.

### **Литература**

- [1]. Вълчанов Хр., И. Юсеинов, Виртуализирана мрежова лаборатория, международна научна конференция UNITECH 2012, Габрово, стр. I473-I477.
- [2]. Иванова, А., А. Смрикаров. Създаване и изследване на виртуална лаборатория по "Организация на компютъра": Дисертация / Анелия Стоянова Иванова, Научен ръководител Ангел Смрикаров. Русе, 2006..
- [3]. Li, P. J. Jones, K. Augustus, Incorporating Virtual Lab Automation Systems in IT Education, American Society for Engineering Education, 2011
- [4]. VMware, VMware Horizon 7 is the leading platform for virtual desktops and applications. <https://www.vmware.com/products/horizon.html>, дата на използване 20.07.2018 г.

#### **За контакти:**

Доц. д-р Делян Георгиев Генков  
катедра „Компютърни системи и технологии“  
Технически университет - Габрово  
E-mail: dgenkov@tugab.bg

# УПРАВЛЕНИЕ НА ЗАХРАНВАНЕТО НА ВИРТУАЛНА ИНФРАСТРУКТУРА

Христо Г. Вълчанов

**Резюме:** Съществен проблем при сървърната виртуализация е отпадането на захранването на клъстер. Повечето UPS системи за виртуализирана инфраструктура се доставят със собствени командни софтуер и интерфейси. Като резултат, администраторите са принудени да използват един инструмент за управление на виртуализацията и друг за управление на захранването, което затруднява производителността и може да забави времето за отговор в случай на отказ на устройство. В настоящия доклад е представена архитектурата на система за управление на захранването и виртуалните машини в инфраструктура, изградена на базата на VMware ESXi клъстер.

**Ключови думи:** Virtualization, UPS, Power Management

## Power Management of a Virtual Infrastructure

Hristo G. Valchanov

**Abstract:** A major problem with server virtualization is the dropping of cluster power. Most UPS systems for virtualized infrastructure are delivered with their own command-line software and interfaces. As a result, administrators are forced to use a virtualization management tool and another for power management, which hinders performance and can slow response times in case of device failure. This paper presents the architecture of a power management system for virtual machines in an infrastructure built on the VMware ESXi cluster.

**Keywords:** Virtualization, UPS, Power Management.

### 1. Увод

Виртуализацията е доказана софтуерна технология, която бързо трансформира ИТ сектора и фундаментално промени начина, по който хората използват компютрите. Виртуализацията все повече се превръща в основа на инициативите, насочени към редуциране на разходите. Според различни проучвания между 60 и 80% от ИТ отделите в средни и големи компании по света реализират проекти за сървърна консолидация. Причините за това са очевидни – чрез редуцирането на броя и видовете сървъри, които осигуряват бизнес приложенията, компаниите редуцират в значителна степен разходите и усилията за управление. Освен това се намалява консумацията на електричество, свързана с работата на сървърите и на охлаждащите системи. Може да се направи заключението, че виртуализацията опростява работата на информационната инфраструктура в компаниите, подобрява производителността и намалява общата цена на притежание на системите. Виртуализационните технологии са в основата на съвременните облачни изчисления [1].

Виртуализационната платформа виртуализира хардуерните ресурси на x86-базирани компютри като създава напълно функционална виртуална машина, която може да изпълнява своя собствена операционна система и приложения, точно като истински компютър [2]. Виртуализацията работи чрез използване на специален софтуер - монитор или хипервайзор, който разпределя хардуерните ресурси динамично и прозрачно между виртуалните машини. Няколко операционни системи (гост системи) могат да работят едновременно на един физически компютър и да споделят хардуерни ресурси един с друг. Чрез капсулиране на цялата машина, включително процесор, памет, операционна система, и мрежови устройства,

виртуалната машина е напълно съвместима с всички стандартни x86 операционни системи, приложения и драйвери.

Съществен проблем при сървърната виртуализация е отпадането на захранването на клъстер. Решението е използване на системи за непрекъсваемо захранване (UPS), които поддържат аварийно захранване от батерии за определен период от време. След достигане на критичен минимум на акумулираната енергия, всички виртуални машини трябва да бъдат коректно изключени (*graceful shutdown*). В противен случай това би довело до евентуална неконсистентност на информацията в тях. Същевременно, след възстановяване на захранването, виртуалните машини трябва да се стартират върху съответните хост-сървъри [3].

Редица компании, като VMware, Microsoft, Citrix и други, предлагат продукти от типа "Live migration", които могат бързо да прехвърлят виртуални машини от един хост-сървър на друг с цел балансиране на натоварването или когато първоначалният сървър има оперативни проблеми или изисква поддръжка. Но предлаганите подобни системи нямат вградена функционалност за реакция при прекъсване на захранването.

Повечето UPS системи за виртуализирана инфраструктура се доставят със собствени командни софтуер и интерфейси. Като резултат, администраторите са принудени да използват един инструмент за управление на виртуализацията и друг за управление на захранването, което затруднява производителността и може да забави времето за отговор в случай на отказ на устройство. При подмяна на системата за захранване е необходимо наново обучение и конфигуриране в зависимост от модела и решението.

Друго предизвикателство е използване на решение, което е преминало достатъчно тестове. Много предлагани решения се рекламират като сертифицирани „готови за виртуализация“, но трябва потребителите да се обръщат към това твърдение с повишено внимание. Реалността е, че много производители не са тествали напълно своя софтуер върху виртуализирани платформи или техните приложения имат някои специфични проблеми във виртуализирана среда.

В настоящия доклад е представена архитектурата на система за управление на захранването и виртуалните машини в инфраструктура, изградена на базата на VMware ESXi клъстер.

## 2. Формулиране на проблема

При разработката на ESXi от компанията VMware не е включен сървис за управление на захранването. Компанията предлага друга технология за защита на техните сървъри. Тъй като за клиентите е недопустимо да си позволят спиране на техните устройства за голям период от време, VMware разработват система, свързана не само с UPS, но и с противопожарна система за безопасност. При активиране в случай на проблем, всички данни от сървърите се криптират, компресират и се прехвърлят на отдалечени сървъри за временно ползване, където се декомпресират и декриптират. Предимствата на този подход се изразяват в това, че нормалното действие на сървърите се прекратява за кратко време и се покриват повече възможни събития, при които е възможна загуба на информация. Недостатък е, че това решение е достъпно само за големи техни клиенти, като е неикономично за малкия и среден бизнес.

Друго решение е предложено от Ралф Лаудерс и е публикувано в официалния форум за отворен код на VMware. Реализацията се базира на възможността на операционната система ESXi да извежда информация за всички физични устройства, свързани към нея. Чрез *cronjob* се проверява всяка секунда за наличие на функциониращи мрежови интерфейси. Ако интерфейс е неактивен за определен период от време, се счита отпадане на захранването на свързания с него комутатор и се преминава към изключване на системата. Подходът е

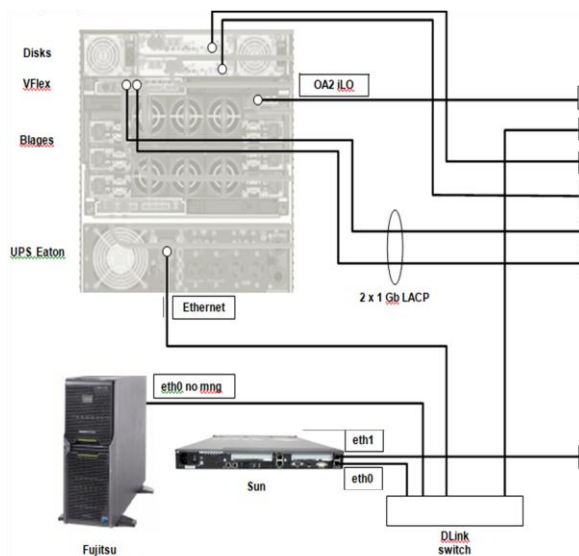


сравнително прост, но не се взема под внимание наличието на UPS и други фактори, като например повишаване на температурата в помещението.

Производителните на UPS, като Eaton, APC и други предлагат собствен софтуер за управление на VMware виртуална инфраструктура. Софтуерът следи състоянието на UPS посредством протокол SNMP и при наличие на изпратено от UPS съобщение се стартира последователност от действия за изключване на машините. Софтуерът използва VMware vCenter Server за комуникация с виртуалните машини [4]. Въпреки удобният интерфейс и възможностите, подобен софтуер има един съществен недостатък – не може да се използва и за коректно изключване на други операционни системи и машини, освен базирани на VMware и ESXi. В редица организации, особено в академичните среди, в комуникационните шкафове често се помещават освен сървърите за виртуалната инфраструктура (като HP Blade) и други машини и устройства с различни операционни системи (Linux или Microsoft сървъри, дискови масиви и др.). Тези системи също изискват да бъдат коректно изключвани.

Налице е и друг важен въпрос, свързан със стартирането на виртуалните машини след възстановяване на захранването. Чрез vSphere клиент-администраторът може да конфигурира автоматичното стартиране на виртуалните машини върху определени хостове, след като самите хостове се стартират. Това дава възможност за балансирано натоварване на сървърите в клъстера. Но една от важните задачи на виртуализационните платформи е осигуряването на наличност и достъпност на приложенията при отказ на оборудването. VMware предоставя услугата High Availability (HA), която позволява ефективно автоматизирано рестартиране за броени минути за всички приложения при срив на хардуера или операционната система. HA постоянно следи всички хостове в клъстера и рестартира виртуалните машини на друг хост, ако оригиналният хост сървър спре да работи. При активиране на тази услуга се изключва възможността за автоматичното стартиране на виртуалните машини върху определени хостове. Вместо това, всички виртуални машини, конфигурирани за автоматично стартиране, могат да бъдат насочени към първия стартиран сървър в клъстера.

На фигура 1 е представена виртуалната инфраструктура, която е необходимо да бъде управлявана. Тази инфраструктура се използва в учебния процес на студентите от специалности „Компютърни системи и технологии” и „Софтуерни и интернет технологии” при ТУ-Варна.



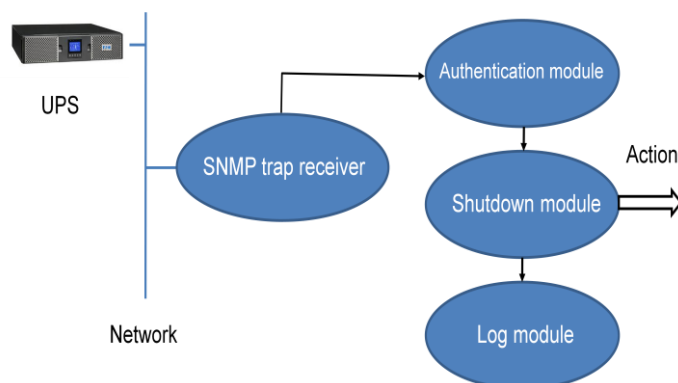
Фиг. 1. Целева инфраструктура

Основният компонент е клъстер от високопроизводителни сървъри HP BL460c Gen8 с VMware виртуализационна платформа, базирана на ESXi 6.0. Върху тях се изпълняват над 50

виртуални машини със специфично предназначение за учебния процес. Виртуалните машини са под Windows и Linux ОС. В комуникационния шкаф са монтирани дисков масив, и две сървърни машини. Захранването се поддържа от непрекъсваемо устройство Eaton 9SX5000.

### 3. Реализация на системата

Архитектурата на системата за управление на захранването се състои от няколко модула (фигура 2).



Фиг. 2. Архитектура на системата

Модулният характер позволява лесно и ефективно модифициране и разширяване на възможностите в бъдеще. Модулите са реализирани като сървиси и скриптове, които се изпълняват на отделен сървър под управлението на Windows Server 2012 R2.

Модулът **SNMP trap receiver** се изпълнява като услуга във фонов режим. Неговото предназначение е да следи в мрежата за наличието на SNMP съобщение с определен код. Това съобщение се генерира от UPS. UPS е конфигуриран да премине в режим на изключване при определен оставащ капацитет на батерията (за целите на изследването това е 65%). След стартиране на процедурата продължителността до изключване на захранващата линия е 600сек. През този период цялата инфраструктура (всички виртуални машини и самите хостове) е необходимо да бъде правилно изключена.

Модулът използва библиотеката за сокети WinSock, като слуша по UDP на порт 162. При получаване на trap съобщение, то се преглежда от специален парсер, който извлича отделните полета на SNMP съобщението. Извършва се проверка за съвпадение на полетата със следната информация:

- Адрес на източника на съобщението - трябва да е IP адресът на UPS.
- OID на UPS – трябва да съответства на идентификатора на производителя на UPS. Идентификаторът се указва посредством конфигурационен файл. Това дава възможност системата да се използва при различни производители на непрекъсваеми токозахранващи устройства.
- Код на generic trap – в конкретния случай е GENTRAP=6.
- Код на специфичен trap – код на начало на процедурата по изключване на захранването на UPS - OFFINPROGRESS=31.

Ако са налице всичките условия, то трябва да се премине към действията по нормалното изключване на инфраструктурата. Стартира се специален скрипт, който изпълнява необходимата последователност от действия. Скриптът е реализиран с използване на PowerShell под Windows и пакета на VMware – VMwareCLI [5,6]. Използването на споменатите технологии е продиктувано от следните обстоятелства. Реализацията чрез скриптове е ефикасен и гъвкав метод за изпълнение на определени действия в операционната



система. Не се изисква използването на компилатор, което дава възможност за бърза модификация при необходимост. Освен това, VMware предоставят мощна библиотека с т.н. *commandlets*, която е интегрирана много добре с PowerShell. Не на последно място, тестването на изпълнението не изисква специални средства, като например дебъгер.

Първоначално се преминава през фазата на автентикация, осъществена от модула **Authentication**. Неговото предназначение е да осигури коректното идентифициране и логване към различните системи.

След успешното свързване се преминава към същинската последователност на изключване. Тя се реализира от **Shutdown** модула, който се изпълнява като скрипт. Действията се различават в зависимост от различните цели, като последователността е следната:

1. Осъществява се свързване към управляващия компонент на виртуализационната платформа – vCenter. Той се изпълнява като отделна виртуална машина върху самостоятелен сървър, различен от клъстера.

```
Connect-VIserver -Server $h.Name -User $BladeUser -Password $BladePassword -WarningAction SilentlyContinue -errorAction "Stop" | Out-null
```

2. След свързване се извлича информация за наличните виртуални машини върху отделните сървъри в клъстера. Всяка изпълняваща се виртуална машина се изключва внимателно при наличие на VMware tools, или принудително, ако не са налични.

```
$esxcli = Get-EsxCli -VMHost $h.Name
# Shutdown the VMs on the host
$vm = Get-VM
Foreach ($vm in $vm) {
    # The VM is powered ON
    if ($vm.PowerState -eq "PoweredOn") {
        $WorldId = ($esxcli.vm.process.list() | where {$_.DisplayName -eq $vm.name}).WorldId

        # This VM has VMware tools running
        if ( ($vm.Guest.Extensiondata.ToolsVersionStatus -eq "guestToolsCurrent") -or
            ($vm.Guest.Extensiondata.ToolsVersionStatus -eq "guestToolsNeedUpgrade") ) {
            $esxcli.vm.process.kill("soft",$WorldId) | Out-null
        }
        else {
            # No VMware tools running - hard stop
            $esxcli.vm.process.kill("hard",$WorldId) | Out-null
        }
    }
}
```

3. След успешното изключване на всички виртуални машини се изключват и самите сървъри в клъстера. За целта е необходимо те да преминат в специален *Maintenance* режим, да се стартира команда за изключване и да се излезе от този режим за нормална бъдеща работа.

```
$esxcli.system.maintenanceMode.set("true") | Out-Null
$esxcli.system.shutdown.poweroff($ShutdownDelay, "") | Out-Null
$esxcli.system.maintenanceMode.set("false") | Out-Null
```

4. Следващата стъпка е изключване на всички останали сървъри. За целта се използва SSH сесия, чрез която се изпълнява команда за изключване на машината.

```
# Prepare credentials
$secpass = ConvertTo-SecureString $SunPassword -AsPlainText -Force
$mycredentials = New-Object System.Management.Automation.PSCredential($SunUser,$secpass)

# Create SSH session with credentials
$session = New-SSHSession $SunIP -Credential $mycredentials

# Run command
$result = Invoke-SSHCommand -SSHSession $session -Command $hostcommand

# Close SSH session
Remove-SSHSession -SSHSession $session | Out-Null
```

5. Последният етап е изключване на самата машина, върху която се изпълнява системата.

При всяко действие се записва информация в журнален файл, която може да бъде използвана впоследствие от администратора за анализ.

#### 4. Заключение

В настоящия доклад е представена архитектурата и базовата функционалност на система за управление на хранването на виртуална инфраструктура, базирана на VMware. Функционалността ѝ не зависи от вида на използваното непрекъсваемо хранване. Системата е модулна, което позволява лесно нейно бъдещо разширяване. Системата е внедрена и в продължение на една година е доказала своята работоспособност.

Като насоки за бъдеща работа се предвижда разработване на графичен потребителски интерфейс, както и интегрирането ѝ като сървис в операционната система ESXi.

#### Литература

- [1]. Rafaels R. Cloud Computing: From Beginning to End. Create Space Independent Publishing Platform. 2015. ISBN 1511404582.
- [2]. Portnoy M. Virtualization Essentials. Sybex. 2012. ISBN 1118176715.
- [3]. Auday Al-Dulaimy, W. Itani, A.Zekri. Power management in virtualized data centers: state of the art. Journal of Cloud Computing Advances, Systems and Applications, 2016.
- [4]. VMware Corporation. <http://www.vmware.com>.
- [5]. PowerShell Documentation. <https://docs.microsoft.com/en-us/powershell/>.
- [6]. VMware PowerCLI Documentation.  
<https://www.vmware.com/support/developer/PowerCLI/>.

#### За контакти:

доц. д-р инж. Христо Г. Вълчанов  
катедра „Компютърни науки и технологии”  
Технически университет - Варна  
E-mail: [hristo@tu-varna.bg](mailto:hristo@tu-varna.bg)



## СЕКЦИЯ 2

# СОФТУЕРНИ И ИНТЕРНЕТ ТЕХНОЛОГИИ

*Пета научна конференция с международно участие  
"Компютърни науки и технологии"  
28 - 29 септември 2018г.  
Варна, България*



*Fifth International Scientific Conference  
Computer Sciences and Engineering  
28 - 29 September, 2018  
Varna, Bulgaria*

## SECTION 2

# SOFTWARE AND INTERNET TECHNOLOGIES

# ERGONOMIC ASPECTS OF CRYPTOCURRENCY MINING

Radosław Wróbel, Celestyn Miłuch, Piotr Haller, Radosław Włostowski

**Abstract:** Extracting cryptocurrencies is now a widely discussed topic. The extraction is not only about having proper infrastructure (equipment, network connections, intermediate devices), but also (and above all) with energy consumption. The article presents the extraction of cryptocurrencies in the light of ergonomics, and more specific - economic. The research was carried out for various scenarios, using various equipment and methods of mining.

**Keywords:** Cryptocurrency, cryptocurrency mining, CPU, GPU

## 1. Introduction

Mining cryptocurrencies (or cryptomining) is a very demanding on hardware as well as software. The process consists in performing compiled mathematical and cryptographic calculations, for which a specific crypto currency is obtained from a chosen team or creator.

The basic mining unit is hash per second (h/s). If our calculations are considered to be correct, they are added to one chain (blockchain) [1].

Cryptocurrency is a distributed accounting system that carries data on a unique key. Key duplication is impossible due to its originality. The advantage of using a virtual currency is a distributed network that exploits the P2P (peer-to-peer) system. As a result, there is no main distribution point, which excludes managing one unit or from one location [2]. A user's wallet is required for storing the currency; this is a private key. It stores selected types of coins. Other wallets may also be required for various types. Wallets in the cloud are a very popular solution. Hardware wallets can also be used. They require connection to the Internet, but can be disconnected at any time from the selected device, thus allowing you to secure your funds without the risk of losing anything via the Internet connection. On the other hand, mechanical damage of such a wallet is possible. If a hardware device is damaged, it is impossible to recover the data stored in it [3].

The existing transactions are confirmed or registered in a chain called blockchain. Data is grouped in blocks and arranged one after the other to form one big string. It contains all the information about purchasing and selling on the web [4]. The data is visible to the public and available to be reviewed by anyone willing to do that. If there is an attempt to change or edit any entry in a block, communication is automatically rejected. The system is so well developed that many sectors of public life are ready to use blockchain to verify and register data. This method allows to preserve the integrity of data, thus counterfeiting is impossible [5].

## 2. Impact of mining on processing

Obtaining virtual currencies is based on many dependent variables. The first is the Return of Investment coefficient (ROI). It determines the profitability and the speed of investment return on the purchased equipment to extract the type of coin we choose. It should be remembered that calculations are carried out for a given day. Profitability may change due to a few parameters. One of them is the 'difficulty' of mining (this parameter determines how difficult it is to extract one block). The higher the coefficient, the fewer coins are mined in the same time; the number of coins depends also on the number of diggers and the price of the currency itself. The price of cryptocurrency is an important factor in mining. The higher the price per coin, the more popular it becomes among miners, as they choose to extract it more often [6]. Another factor is high electricity

consumption by our mining equipment. It is essential that our mining set should have the greatest possible computing power, and, at the same time, consume as little electricity as possible.

The first mining test will be performed with the use of a web browser. During the working process, only the main computer processor (CPU) is used. The tests were carried out in the 'normal' (office) use of the device by the owner. Depending on how many CPU cores there are and how large their computing power is, you can achieve better or worse results in mining. The test was carried out on an Intel i7 processor with eight cores and each with 2.5 GHz tact. The tests were carried out using one core, two cores, then four, six and eight cores.

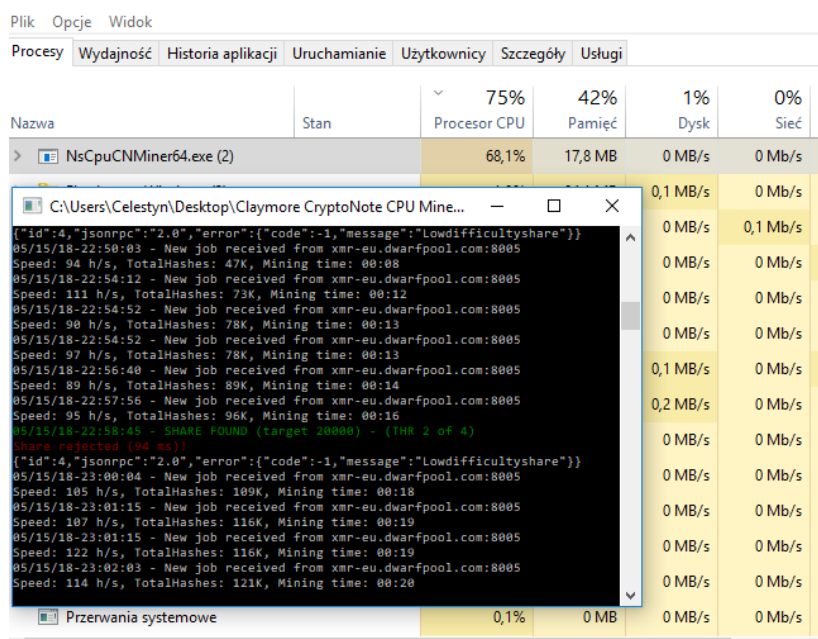
Using one core, the CPU usage was noted at 25%. The CPU usage is so high that more advanced users already suspect the presence of software from the outside. The use of the internet connection was noticed only at the start of the application, when the connection to the server was established. The cryptocurrency extraction itself did not achieve significant results. Achieving the extraction at the level of 13.5 hashes per second is very ineffective. The power consumption of the entire device makes it very unprofitable.

The following table presents the results of subsequent tests. None of them became profitable for a single user despite the increase in mining speed.

**Table 1.** Velocity of cryptocurrency mining vs. number of cores (processing)

Number of cores	Extraction speed	Use of processor
Two cores	21.8 h/s	47%
Four cores	24.3 h/s	65%
Six cores	39-40 h/s	90-95%
Eight Cores	48 h/s	100%

The second test was carried out directly on a device dedicated to mining. An extraction script had been installed for the cryptocurrency, which uses all (if possible) computing power of the CPU. The same processor was used in the test as in the previous one. In this case, the full power of the processor was used immediately, but with different use from the device. With normal use, the production level at 94-117 h/ s was reached. Operating the computer was very onerous. All processes were very slow and required a lot of patience of the user, as it took a relatively long time to execute commands. In the Task Manager, less CPU use is noticeable. When using a web browser at maximum efficiency, the processor was used at 95–100%. Here, the average computer component is used at 75%.



**Fig. 1.** Script responsible for extraction and processing

Another test was carried out under ‘little-collision’ conditions (without the owner using the computer whatsoever). Only the mining program was running on the computer. This allowed for excluding external factors, which in part may have compromised the results of the previous test. Under ideal conditions, a significant increase in extracting efficiency was noted at 159 h/ s. Unfortunately, the above methods for a single unit are economically unprofitable. With an average yield of 105 and 159 h/ s, you can only make a loss.

During the digging, a session was taken over (link sniffing). High TLSv1.2 traffic is generated, which is used to transfer information between the server and a virtual digger. Frequent propagating of GQUIC (Google Quick UDP Internet Connections) protocol is also noticed. This protocol is used for communicating in the services such as YouTube or Facebook. Its task is to shorten the time of connecting the host with the server. It is not related to the extraction of cryptocurrency at all.

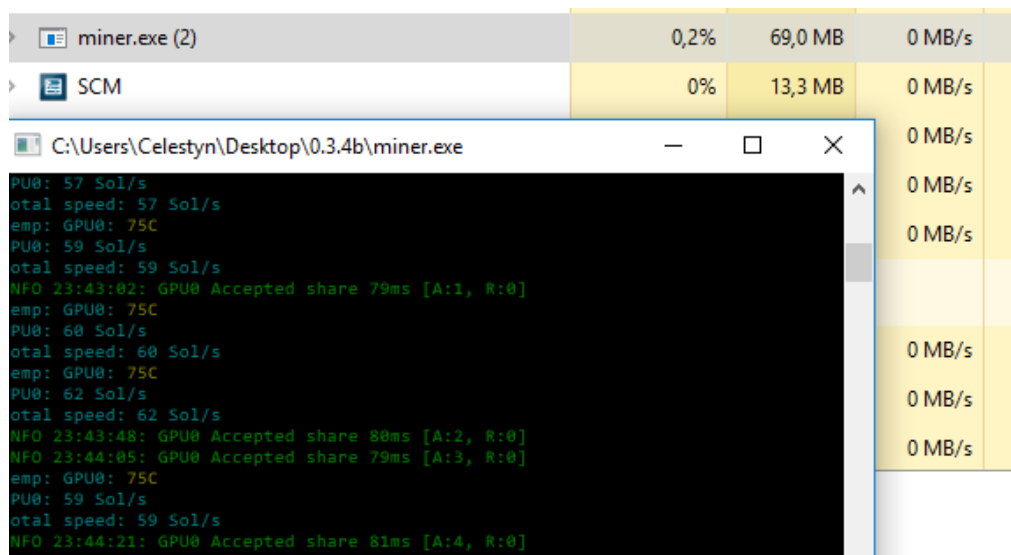
24	5.326005	209.58.153.115	192.168.1.12	TLSv1.2	16/ Application Data
25	5.367382	192.168.1.12	209.58.153.115	TCP	54 7554 → 443 [ACK] Seq=1 Ack=437 Win=251 Len=0
26	5.421038	209.58.153.115	192.168.1.12	TLSv1.2	98 Application Data
27	5.421641	192.168.1.12	209.58.153.115	TLSv1.2	102 Application Data
28	5.557986	92.38.169.14	192.168.1.12	TLSv1.2	85 Application Data
29	5.558337	192.168.1.12	92.38.169.14	TLSv1.2	89 Application Data
30	5.586528	209.58.153.115	192.168.1.12	TCP	54 443 → 7551 [ACK] Seq=45 Ack=49 Win=76 Len=0
31	5.691561	92.38.169.14	192.168.1.12	TCP	54 443 → 7552 [ACK] Seq=32 Ack=36 Win=63 Len=0
32	5.715507	209.58.153.115	192.168.1.12	TLSv1.2	168 Application Data
33	5.757125	192.168.1.12	209.58.153.115	TCP	54 7554 → 443 [ACK] Seq=1 Ack=551 Win=251 Len=0
34	5.940871	209.58.153.115	192.168.1.12	TLSv1.2	98 Application Data
35	5.941582	192.168.1.12	209.58.153.115	TLSv1.2	102 Application Data
36	6.106333	209.58.153.115	192.168.1.12	TCP	54 443 → 7554 [ACK] Seq=595 Ack=49 Win=123 Len=0
37	6.325511	SamsungE_8a:40:0b	Broadcast	ARP	60 Who has 192.168.1.1? Tell 192.168.1.10
38	6.653874	SamsungE_8c:a8:a4	Broadcast	ARP	60 Who has 192.168.1.1? Tell 192.168.1.11
39	6.935363	209.58.153.115	192.168.1.12	TLSv1.2	149 Application Data
40	6.976328	192.168.1.12	209.58.153.115	TCP	54 7554 → 443 [ACK] Seq=49 Ack=690 Win=256 Len=0
41	7.143018	209.58.153.115	192.168.1.12	TLSv1.2	168 Application Data
42	7.184198	192.168.1.12	209.58.153.115	TCP	54 7554 → 443 [ACK] Seq=49 Ack=804 Win=256 Len=0
43	8.093404	192.168.1.11	239.255.255.250	UDP	77 38545 → 15600 Len=35
44	8.341544	SamsungE_8a:40:0b	Broadcast	ARP	60 Who has 192.168.1.1? Tell 192.168.1.10
45	8.673584	SamsungE_8c:a8:a4	Broadcast	ARP	60 Who has 192.168.1.1? Tell 192.168.1.11
46	9.839772	209.58.153.115	192.168.1.12	TLSv1.2	168 Application Data
47	9.880993	192.168.1.12	209.58.153.115	TCP	54 7554 → 443 [ACK] Seq=49 Ack=918 Win=255 Len=0
48	10.272481	HuaweiTe_f4:0b:7e	Broadcast	ARP	42 Who has 192.168.1.1? Tell 192.168.1.21
49	10.365605	SamsungE_8a:40:0b	Broadcast	ARP	60 Who has 192.168.1.1? Tell 192.168.1.10
50	10.697727	SamsungE_8c:a8:a4	Broadcast	ARP	60 Who has 192.168.1.1? Tell 192.168.1.11
51	10.954517	209.58.153.115	192.168.1.12	TLSv1.2	168 Application Data
52	10.996364	192.168.1.12	209.58.153.115	TCP	54 7554 → 443 [ACK] Seq=49 Ack=1032 Win=255 Len=0
53	11.093727	192.168.1.11	192.168.1.255	UDP	77 51388 → 15600 Len=35

**Fig. 2.** Communication session during extraction of cryptocurrency

The program for extracting cryptocurrencies has not been detected by the security systems as ‘harmful.’ In addition, the antivirus program does not interpret the behaviour of the script executing file as malicious. It is interpreted in this way only when the script is run manually. This makes it difficult to diagnose the cause of increased wear of components even when the computer is idle.

The last test to be carried out was the extraction of cryptocurrency with the use of a graphics card. This is the most popular method among ‘virtual miners.’ A large number of efficient graphics cards allow for efficient ‘extraction’ of popular virtual currencies. A card dedicated to mobile devices from the NVidia Geforce family was used in the test. After starting the program, the temperature of the graphics card was noticeably higher. After a few seconds, additional cooling was initiated. The extraction itself on the NVidia board in the desktop version was not large. The mining numbers ranged from 57 to 65 solutions per second. This result is below profitability. Compared to graphics cards dedicated to desktops, this performance was five times as low.

The mining program itself, apart from the temperature increase in the heatsinks, does not show any other ‘symptoms.’ In the previous test, it was possible to check in the Task Manager which units are used. In this case, the program does not show any other signs of wear. However, keep in mind that if you begin to suspect any operation of external software, you can check it in a different way.



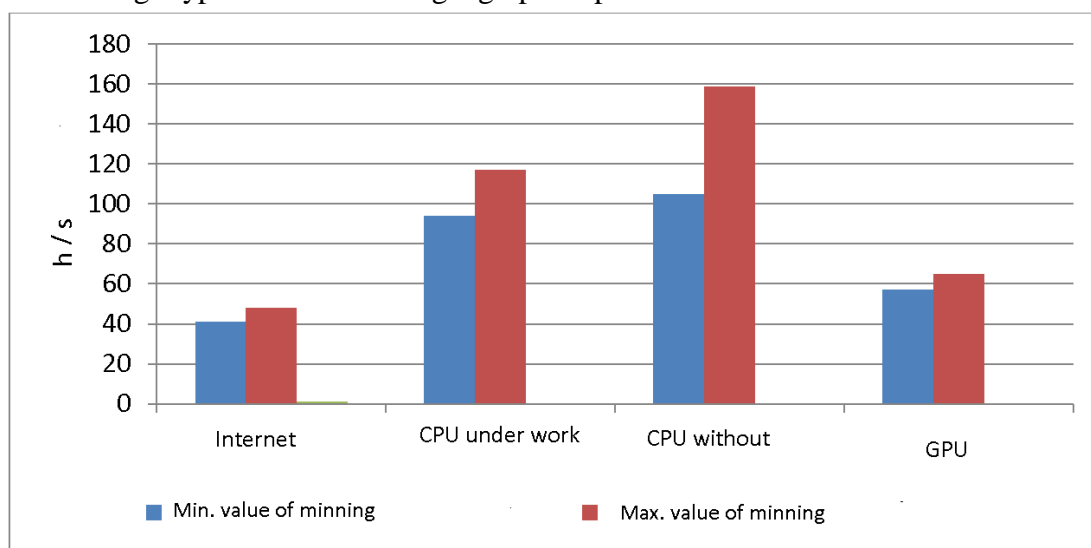
**Fig. 3.** Script responsible for extraction and processing when a graphics card is used for mining

For mining to be effective, the software must be running all the time. Therefore, the basic functionality of such programs is autostart. They are there because of the maximum use of the graphics card for calculations. In order not to arouse suspicion, starting is set to a delayed start. This allows us to avoid collision with the starting system and to use the component without any problems at its full power. It does not cause suspicion on the part of the user who is not aware of the intruder, and is used by third parties in a way that is not fair.

### 3. Conclusion

The research objective was to test the performance of the computer while extracting selected cryptocurrencies. Initial tests were carried out, which allowed the experimenters to select appropriate tools as well as the best hardware configuration. Then, performance tests were performed. The chart below summarizes the most important tests:

1. Digging using a browser with full load-bearing power.
2. Digging using the main processor of the computer at initiated extraction.
3. Coin extraction using a CPU under ideal conditions.
4. Extracting cryptocurrencies using a graphics processor.



**Fig. 4.** Summary of tests



The above tests (Fig. 4) show that digging can be done without being noticed by an average user. All transmitted data is fully protected by encrypting a session. In the case of web browsers, plugins of unknown origin can be dangerous. Distributing such plugins by their creators in the Internet gives the prospect of a large income with little effort. The flow of data sent from the computer to the server has been checked. The next tests were to show what threats are associated with unauthorized installation of mining software. Thanks to a powerful main processor, it was possible to check how the device behaves when particular units are gradually loaded. It was also possible to determine how the communication works and what data packages are employed in the session. Research using a graphics card showed that despite the invisible consumption of network resources, the device can also participate in the mining process. The data transfer between the end devices itself did not show excessive use of the network connection. The most characteristic symptom was the increased temperature of the graphics processor. It should be remembered that the result obtained with the use of the graphics card depends on the version of the drivers and the model of the graphics card itself. The cards dedicated to desktops and recognized as the leading products in their class will be much more efficient.

All tests were carried out under normal conditions of use. This allowed to simulate the use of a computer by an ordinary user. For people with little knowledge in the analysis of network packets, the problem may not be noticed.

### References

- [1] Description of the Hashrate concept <https://ceo.com.pl/hashrate-bonment-processing-processing-secretaryuture-26967>.
- [2] Description of the concept of Cryptocurrency <https://businessinsider.com.pl/finanse/kryptowaluty/czego-sa-bitcoiny-i-kryptowaluty/05f5ned>.
- [3] Comparison of cryptocurrency wallets <http://traderoom.pl/poradniki/portfele-kryptowalut>.
- [4] Rudiments of Blockchain <https://norbertbiedrzycki.pl/blockchain-trzeba-o-nim-wiedziec>.
- [5] Blockchain technology description <http://kryptopolonia.info/co-to-jest-lancuch-blokow-blockchain>.
- [6] Description of the ROI concept <https://blog.hashflare.io/crypto-mining-returns-what-should-you-consider-78e6c2ef0a23>.

### For contacts:

DSc Radoslaw Wrobel  
Division of Vehicle Eng.  
Wroclaw University of Science and Technology  
E-mail: [radoslaw.wrobel@pwr.edu.pl](mailto:radoslaw.wrobel@pwr.edu.pl)





# TOWARDS A PERSONALIZED AND ADAPTIVE E-LEARNING PLATFORM

Barış Altop, Mehmet Cem Aytekin, Muhammed Mucahid Benlioğlu, Stefan Rabiger, İnanç Arın, Duygu Karaoğlu Altop, Yücel Saygın

**Abstract:** The amount of people taking online courses has skyrocketed in the recent years. Additionally, starting from universities, most schools have gone under a digital revolution and adapted online e-learning tools to convey information and resources to the students and ease the work of instructors. Although such new and ground-breaking tools have been introduced to the education system, the main problem is still at loose, where all learners are given the same information in the same manner and expected with the same outcome. On the contrary, people tend to have different learning curves or perform better when information is conveyed in various other methods, expecting a more personalized experience. Our platform aims to fill this gap by tracking learner achievements and by introducing the micro-topic concept.

**Keywords:** Learning technologies, e-learning platform, adaptive learning, personalized learning, pervasive learning, active learning, self-paced learning

## 1. Introduction

In traditional tutor-based learning systems, instructors take into account the fact that no two learners are the same of each other by means of strengths and weaknesses [1]. Each and every person has different capabilities and requirements. Therefore, the speed of the learners and the rate at which they learn will always be different. This brings out the importance of the learning preferences. Some people may learn better when they listen something, some learn better with visualization and some learn better with simply reading the content. Thus, modern learning systems need to imitate the role of the instructors of the tutor-based learning systems to provide efficient, personalized and adaptive learning experience. This can be accomplished through having mastery learning that inholds adaptive personalized learning paths rather than having static learning sequences for the contents to be delivered. Learners should be able to follow the contents at their own pace, in their own place and with their own convenience.

With the advancements in e-learning technologies, classroom based education give place to the tools that extinguish the geographical limitations. Another important advancement that stems from this technology is that we can now provide instant feedback to the learner, which makes it possible to have active learning with self pace.

In the fast consumption world that we are living in now, everything is changing very rapidly, so as our attention span. In such a world, adaptability in learning technologies has the utmost importance. Both the content and the delivery mechanisms should keep pace with the requirements of the new age. In this regard, with this paper, we outline an adaptive pervasive learning framework using which instructors can provide the content in the best possible way with the latest technology, and the learners can improve themselves by following the best possible path in their learning journey. Learners using this proposed framework can have more intelligent and personalized learning experiences as provided by the underlying adaptive model.

## 2. Related Work and Problem Definition

Adaptive learning methodologies have been debated since the late 1980's and has evolved in many aspects with the introduction and advancements in internet and computer science, respectively. Since the dawn of the topic, many models and ideas have been applied and has

incrementally made the field of research deeper. With the addition of modern artificial intelligence methods, learning analytics *can* and *has* become more personal for the learner.

In 1990's, hypermedia based systems [2] started to emerge and before the end of the decade, both adaptivity and online techniques made its way into the literature [3-6]. Once online learning platforms became popular, the topic has shifted to creating better and more personalized learning management systems (LMS) for all students, instructors, companies and educational institutions [7].

A uniform and straight learning model is no longer a good fit for all learners with different backgrounds and knowledge levels, hence more content for various levels have been created. Many have tried to use this newly available content to create better learning models for different people rather than personalizing it as if all people have the same capabilities and requirements [8-9].

The problem has been attacked from various perspectives, where the learners were evaluated: (i) within the course itself, (ii) based on their overall performance within the system regardless of the topic or field, (iii) whether or not they should be presented with hints, (iv) whether or not one should repeat the section.

Although promising results have been presented, a complete and unified solution to the problem has not been developed. Most of the models have been successful in solving singular aspects of the problem of an adaptive pervasive learning system, however they all have taken the content of any course as "*as-it-is*" [10-12].

In contrast to previous works, the model this paper proposes offers a new insight to the problem: here, we present the notion of micro-topic for the courses in order to better analyze and adapt the needs of the students. In addition to that, learner is profiled not only based on his/her previous success in all courses taken and his/her current level in an active course, but also on his/her preferences for the courses.

### 3. Our Proposed Adaptive Pervasive Learning Platform

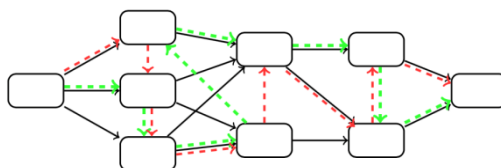
Since no two people are the same, one cannot expect them to learn a topic or a course in the same way. Some instructors in smaller classrooms have the advantage to capture such differences and adapt their teaching skills to maximize the learning curve of each person. Our model tries to understand this procedure and implement it to the benefit of both the instructor and the learner.

Our model consists of two parts, which complement each other. During our research, we have come to the conclusion that by just introducing adaptivity to a course or a set of courses does not reach the potential of its sophistication. Hence, we took a broader look into both the courses and the underlying adaptivity that could be introduced. We have reached to the understanding that the delivery of the course content should be modified as well.

In the subsections given below, we introduce the components of our adaptive pervasive learning system that builds upon these initial observations.

#### 3.1. The Notion of Micro-Topic

Our proposed adaptive pervasive learning framework (APL) is built upon a X Graph (XGraph), which is a directed acyclic graph as given in Fig. 1. The nodes of this XGraph are called as micro-topics, and the edges between the nodes define the prerequisite relationships among the micro-topics.



**Fig. 1.** XGraph - Green and red dashed lines show the possible paths of different users of the given XGraph

A micro-topic is a single, tightly defined content that is delivered to the learner in different kinds of formats such as videos, animations, slide shows, simple text, etc. In other words, it is the smallest unit in which there is the instructional content. Each and every micro-topic has a set of prerequisites and a set of learning outcomes. For the learner to be able to start with a particular micro-topic, (s)he has to satisfy all of the prerequisites of the target micro-topic. Therefore, the prerequisite set of the target micro-topic must be a subset of the union of the learning outcomes of the respective micro-topics.

Besides, within a micro-topic, there exists a set of resources through which the particular content is delivered. Different instructors can teach a specific content in different ways, or even, a particular instructor can teach a specific content in many different ways. Learners also differ in the sense that some may grasp the idea better if they listen the content and the others may go for reading. In order to handle such diversities, each micro-topic is composed of content deliveries at different levels and by means of different formats. In this regard, a resource belonging to a micro-topic has 3 attributes: (1) format, (2) instructor, and (3) level. As the names imply, format is the type of the content delivery, i.e. video, text, etc., and the instructor is the information regarding the person delivering the content. On the other hand, level specifies the level of teaching that is congruent with the learning curve of the learner.

When the learner is to study a specific micro-topic on XGraph, the Adaptive Path Finder (APF) selects a resource from the respective micro-topic set based on the learner's profile. The knowledge adequacy of the learner on the learning outcomes of the studied micro-topics are evaluated through assessments and the interaction of the learner with the proposed APF framework constructs his/her learner profile. After the selected resource is delivered to the learner, an assessment can be carried out to determine to what extent the learner can satisfy the learning outcomes of the specific micro-topic, where assessments are sub-nodes of micro-topics and each micro-topic can have multiple assessments and can be carried out depending on the learners expected learning curve. If the learner has enough knowledge to continue with the upcoming micro-topic, the same procedure described above will be carried out for the target micro-topic. Otherwise, depending on the evaluation of the assessment, the APF will direct the learner to another resource from the respective micro-topic set. Therefore, at each micro-topic content delivery, a decision is made on the best possible resource for the learner to study next so that (s)he can master the respective micro-topic.

### **3.2. The Adaptive Path Finder**

Behind the XGraph architecture lies a powerful artificial intelligence (AI) system, that will be working in real-time gathering learner data and assessing them into a unique profile for each course they have taken and are actively taking. We call this system the Adaptive Path Finder (APF). The AI module will follow the learner's assessments from each micro-topic in the XGraph and will compare it with its historical data in order to better understand the learner's current state and to offer the best path for a comfortable success in the course. As stated in Section 2, the learners' success degree in various fields will also be accounted for during this profiling process.

The AI module will be profiling a learner from the moment (s)he is introduced to the system. At first, it will rely on other learners' data and will try to match the new learner to one of them as the initial learning data of the new person will be scarce. With existing learners, new learners and/or new courses with great amount of new micro-topics, the data flow will increase exponentially, making AI module better in analyzing and profiling. The initial goal of the AI module will be to correctly assess learner performance and guide the learner through a path that will ensure the success of the learner in the course.

The profile will contain the learner's natural abilities to certain fields, his/her past learning paths with respect to the course fields, the previous predictions made by the system about the

learner and the accuracy of those predictions. It will also take over the curriculum for the learner and the course-offerer. This way, the offerers will not be burdened under the task of editing the curriculum in order to reach out to every learners' needs and the learners will not be forced to undergo with the following the curriculum while trying to understand the content.

The APF will be built from several components, each of which are described in the subsequent subsections. Each one of these components will take over a specific task within the engine and will contain models for several machine learning algorithms for enforcing better adaptivity. The entire engine will work as an AI tool, which will guide each learner personally among each course they take.

### **3.3. Learner Profiling**

Up until recently, a learner's success on a set of subjects were determined by a single exam resulting in a number. The disadvantage of this approach is that the exam performance of a learner depends on multiple factors and can cause big variations in the result that the learner gets from the exam. Hence, this does not completely measure the learner's ability towards the topics covered by the exam.

Our proposed model will incorporate both the new micro-topic concept and the classical assessment model. After each micro-topic is delivered, the knowledge gain of the learner will be assessed through a series of questions. It will then be combined with the data collected during the time the learner has spent on the contents, and a score will be calculated based on the previous performance of the learner and the current assessment results using probabilistic machine learning models.

Our proposed model will also be capable of understanding the learner's skills and capabilities to different topics by combining all the assessments for each course the learner takes. The system can then make the decision on whether the learner is more skilled in natural sciences or in arts and social sciences. Depending on this knowledge, the system will guide the learners through different levels of micro-topics and different paths, even if they are taking a course in contrast to their skills and capabilities.

### **3.4. Content, Resource and Knowledge Profiling**

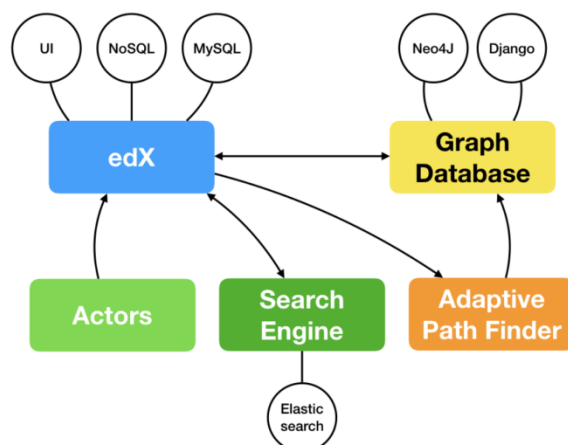
Micro-topics of the same content will contain the same information in different formats and in various levels. This will give the system to offer the best content for each individual's learning habits. The format of each micro-topic can be, but not limited to, either in one or many of these formats: (i) Text, (ii) Slides, (iii) Video/Animation, and (iv) Code (Programming Language). In addition to different content formats, each micro-topic can consist of different levels of difficulty depending on the learners skills and learning preferences. Depending on the learner's profile, the system will decide which types of contents suit a learner best.

Besides contents having levels for each learner, the system will use the learner profiles and the content profiles to find the resources and their types that will fit the learners abilities best. The idea behind this method is that, when an instructor decides to help out a learner, it will first assess the knowledge of the learner and then will try to convey the new knowledge in different means. Our system will follow on this method and give the learner, for each micro-topic, the best resource type that will have the most positive impact on the learners success on that topic.

By monitoring the learner during the courses that (s)he takes and measuring his/her overall knowledge, the system can infer on the learner's knowledge at any given time. Through this decision process, the system can evaluate the learner, before even (s)he starts a new course.

### 3.5. General Architecture

Our proposed model consists of five building blocks: Actors, e-learning platform, graph database, search tool and artificial intelligence tool APF. The relationship among these blocks are shown in Fig. 2.



**Fig. 2.** Building Blocks - The five major blocks and their underlying technologies.

The actors of the system are divided into three categories: (i) Learners, (ii) Instructors, and (iii) System Administrators. The main idea of this division is to maintain the access rights for content and control over the content. Learners are the base user, where everyone can be a learner no matter their status within the system. Learners can take courses that is offered to them and follow their process, while advancing in the course by following the micro-topics. Instructors are the actors that can create new micro-topics or add new alternatives to any micro-topic. They can also create entire courses by using both already-available micro-topics and their own content that they can add as a micro-topic. On the other hand, system administrators have control over the entire system. They can create and delete users as well as add or remove roles to them. They can also create and control courses and assign courses to learners or grant learners access to courses depending on the control levels they might have.

We have chosen to use an e-learning platform in order cut development time and costs. Using an openly available system gives us the ability to just develop the necessary functionality we see fit for our use over the existing back-end of the platform. We have chosen to use Open edX [13] as our platform, which is capable of creating and maintaining the model we have proposed. Open Edx has the ability to present various forms of content in a user-friendly browser interface.

In order to contain courses and their relationships we have decided to use a graph database. The advantage of using such a database is primarily speed and adaptivity by avoiding costly functions like joining, which is replaced by edge and node relationships. The nodes in the graph database will represent the micro-topics of the course: It will store important information about the micro-topics, however it will not store the actual content because the content will be available and programmed as provided by Open edX. The nodes will also store other features that will define the type of the topics for the APF Engine, such as tags, prerequisites, post-requisites, etc. On the other hand, the directed edges of the graph represent the most basic prerequisite relationship for a course-taker between two micro-topics. Each micro-topic may lead to different micro-topics depending on the course and the assessment of the course-taker. The prediction of a course-takers level can be adjusted at various points during the course. This will lead to variations in the edges destination node.

Although both the e-learning platform and our graph database have separate search abilities, we have concluded that when the size of the data the system produces, increases a more powerful search tool will be needed. Hence, we have decided to use elastic search as our main search

framework, where it will traverse the database and build its own index for faster and more precise queries.

As detailed in Section 3.2, the adaptive path finder (APF) system is the key ingredient in our model, in which the tool will be constantly monitoring learner actions and find out what path during the course will benefit his/her learning process at best. Learner actions such as time spent on micro-topic, time spent on assessment, how many times the learner has interacted with the content, e.g. pausing or replaying parts of a video, going back during a slideshow, etc. will be monitored and counted as data points by the APF and will be compared to previous knowledge on the learner. With each micro-topic, the APF will be able to build a more accurate profile of the learner.

#### 4. Conclusions

We have proposed an adaptive online learning platform model, that can learn from both learners and content and adapt content in favor of the learners' abilities to maximize his/her understanding of the topics. In order to achieve high level precision, we have also introduced the concepts of micro-topic and XGraph.

Compared to traditional course layout, consisting of chapters and sections, the micro-topic contains byte-sized information with different presentation types trying to convey the information fast and clear to the learner.

Instead of using a classical table-based database management system in order to store data we have chosen to use a graph database tailored to fit our micro-topic model, which we call the XGraph. The XGraph is highly versatile and can evolve easily into new models.

Finally, the APF collects and analyzes the data provided by the learner, micro-topic content usage and XGraph models, and uses the results to profile the learners, the contents and the relationship between them. Through this the APF is able to make on-the-fly decisions as the learner advances in the courses he/she is taking simultaneously and independently.

We have been working on creating content for an introductory level programming course and will start testing the micro-topic concept on approximately 650 students. During these test runs the initial training data for the APF will be collected and will be used for the following semester as the first version of the system is presented to the students.

#### References

- [1]. Richard M. Felder and Rebecca Brent. Understanding student differences. *The Research Journal for Engineering Education*, 94(1):57-71, 2015.
- [2]. Taekyong Lee, Lei Sheng, Tolga Bozkaya, N Hurkan Balkir, Z Meral Ozsoyoglu, and Gultekin Ozsoyoglu. Querying multimedia presentations based on content. *IEEE Transactions on Knowledge and Data Engineering*, 11(3):361–385, 1999.
- [3]. P Brusilovsky, E Schwarz, and G Weber. A tool for developing hypermedia-based its on www. In *Proceedings of Workshop" Architectures and Methods for designing Cost-Effective and Reusable ITs"* at the Third International Conference on Intelligent Tutoring Systems, ITS, volume 96, pages 12–14, 1996.
- [4]. Peter Brusilovsky and David W Cooper. Adapts: Adaptive hypermedia for a web-based performance support system. In *Proceedings of the 2nd Workshop on Adaptive Systems and User Modeling on the WWW*, volume 41, page 47, 1999.
- [5]. Peter Brusilovsky et al. Adaptive educational systems on the worldwide-web: A review of available technologies. In *Proceedings of Workshop" WWW-Based Tutoring"* at 4th International Conference on Intelligent Tutoring Systems (ITS'98), San Antonio, TX, 1998.
- [6]. John Eklund and Peter Brusilovsky. The value of adaptivity in hypermedia learning environments: A short review of empirical evidence. In *Proceedings of Second Adaptive Hypertext*

and Hypermedia Workshop at the Ninth ACM International Hypertext Conference Hypertext, volume 98, pages 11–17, 1998.

- [7]. John S Kinnebrew, James R Segedy, and Gautam Biswas. Integrating model-driven and data-driven techniques for analyzing learning behaviors in open-ended learning environments. *IEEE Transactions on Learning Technologies*, 10(2):140–153, 2017.
- [8]. Rianne Conijn, Chris Snijders, Ad Kleingeld, and Uwe Matzat. Predicting student performance from lms data: A comparison of 17 blended courses using moodle lms. *IEEE Transactions on Learning Technologies*, 10(1):17–29, 2017.
- [9]. Maren Scheffel, Hendrik Drachsler, Joop De Kraker, Karel Kreijns, Aad Sloodmaker, and Marcus Specht. Widget, widget on the wall, am i performing well at all? *IEEE Transactions on Learning Technologies*, 10(1):42–52, 2017.
- [10]. Colm Howlin and Danny Lynch. Learning and academic analytics in the realizeit system. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, pages 862–872. Association for the Advancement of Computing in Education (AACE), 2014.
- [11]. Com Howlin and Danny Lynch. A framework for the delivery of personalized adaptive content. In *Web and Open Access to Learning (ICWOAL)*, 2014 International Conference on, pages 1–5. IEEE, 2014.
- [12]. Danny Lynch and Colm P Howlin. Real world usage of an adaptive testing algorithm to uncover latent knowledge. In *7th international conference of education, research and innovation (ICERI2014 proceedings)*. IATED, Seville, Spain, pages 504–511, 2014.
- [13]. edX Inc. Open edX Portal. <https://open.edx.org> (2017).

**For contacts:**

Prof. Yücel Saygın  
Faculty of Natural Sciences and Engineering  
Sabancı University  
E-mail: [ysaygin@sabanciuniv.edu](mailto:ysaygin@sabanciuniv.edu)



# МЕТОДИ ЗА ИЗСЛЕДВАНЕ НА СЛУЧАЙНОСТТА НА НЕДВОИЧНИ ПСЕВДОСЛУЧАЙНИ ПОСЛЕДОВАТЕЛНОСТИ

Антония Т. Ташева, Боян Б. Петров, Камен Н. Стойков

**Резюме:** С развитието на компютърните технологии нараства и приложението на псевдослучайните последователности, а от там и нуждата от създаването на нови генератори на такива последователности. Те трябва да бъдат обстойно тествани, за да се определи характеризиращото ги ниво на случайност и съответно възможното им приложение в практиката. Статията разглежда съществуващите софтуерни инструменти като NIST STS, Dieharder и TestU01 и предлага методология за тестване на недвоични псевдослучайни последователности. Предложени са допълнителни тестове, които следва да бъдат проведени.

**Ключови думи:** псевдослучайни последователности, NIST, DieHard, DieHarder, ENT, TestU01, LIL-базирани тестове, ASIN-базирани тестове

## Methods for Randomness Evaluation of Nonbinary Pseudorandom Sequences

Antoniya T. Tasheva, Boyan B. Petrov, Kamen N. Stoykov

**Abstract:** As the application of pseudorandom sequences grows with the development of the computer technologies, the need of new pseudorandom generators increases. Such propositions should be widely tested for their randomness characteristics in order to justify their application. The paper examines known software utilities like NIST STS, Dieharder and TestU01 and proposes a methodology for testing nonbinary pseudorandom sequences. Some additional tests that should be conducted are stated.

**Keywords:** pseudorandom sequences, NIST, DieHard, DieHarder, ENT, TestU01, LIL-based tests, ASIN-based tests

### 1. Увод

С развитието на компютърните технологии нуждата от случайни числа неимоверно нараства. Последователностите от случайни числа намират широко приложение в криптографията, симулации чрез Монте Карло методите, електронни и хазартни игри. Нарастващата нужда от голямо количество случайни числа е причината за появата на така наречените генератори на псевдослучайни числа (Pseudo Random Number Generators – PRNG). Генераторите изпълняват поредица от аритметически операции, в резултат на които се определят числа, наподобяващи случайни. Поради детерминистичната същност на генераторите изходните им последователности не могат да бъдат истински случайни и се наричат псевдослучайни. През последния век са предложени и изследвани множество генератори, най-често основаващи се на линейни преобразувания или рекурсивни функции.

И до ден днешен продължава търсенето на генератори, осигуряващи голям период на повторение, добри статистически свойства, ниска автокорелация и други изисквания, наложени от сферата на приложение. Така в практиката са се наложили различни методи за оценка дали един генератор на псевдослучайни числа е достатъчно „добър“ или „лош“.

Тестването дали една последователност е наистина случайна следва да бъде безкрайно, съгласно определението за случайност. Разбира се, това е неприложимо на практика. Така следва да се определят нива на случайност, за които да се проверява, в зависимост от предвиденото приложение на генерираните последователности.

Съществуват математически доказателства за определяне на периода, линейната сложност и други емпирични показатели на генераторите, докато в същото време



изчислението на постигнатата случайност не е толкова лесно за измерване. Нуждата от средства за оценка на случайността обуславя разработката на методи и в следствие създаването на софтуер за автоматизирано извършване на тестове. Така през 1995 година George Marsaglia публикува тестовия пакет DieHard [1, 5], който дълго време е най-използваният. По-късно се появяват и налагат препоръките на националния институт за стандарти и технологии (National Institute of Standards and Technology – NIST) [6] и съответният им пакет тестове. В статията ще разгледаме и набора от тестове TestU01 [4], предложен от Pierre l'Ecuyer и Richard Simard, както и представените последните години тестове, базирани на закона на повторения логаритъм (Law of the Iterated Logarithm – LIL) [12], както и статистически тестове, базирани на аркус синус закона (Arcsine Law – ASIN) [7].

Статията е структурирана както следва. Първо са разгледани съществуващите софтуерни инструменти за оценка на случайността на изходните последователности на генератори на псевдослучайни числа. Обобщени са включените в тях тестове. Предложена е методология за тестване на недвоични псевдослучайни последователности. Разгледани са допълнителни тестове за оценка.

## 2. Обзор на съществуващите инструменти за оценка на двоични последователности

**Статистическите NIST тестове** [8, 9] проверяват валидността на специфичната нулева хипотеза  $H_0$ : “Тестваната последователност е случайна”. При това алтернативната хипотеза  $H_1$  гласи, че последователността не е случайна. За всеки от включените 15 теста се взема решение, което приема или отхвърля  $H_0$ , т.е. определя се дали PRNG генерира или не случайни стойности.

За всеки тест се изчислява статистическа *P-стойност* като функция на изследваните данни. За всички 15 на брой NIST тестове, *P-стойността* [8] се дефинира като „вероятността за това идеалния случаен генератор на числа да генерира последователност по-малко случайна от тестваната последователност”. Така дефинираната *P-стойност* показва доколко изследваната последователност е случайна. В частност, ако *P-стойност* = 1, то изследваната последователност е абсолютно случайна. От друга страна, нулева *P-стойност* индикира напълно неслучайна последователност.

**Таблица 1.** Възможни заключения от статистическите тестове

Реална ситуация	Заключения	
	Приема се $H_0$	Приема се $H_1$
Последователността е случайна ( $H_0$ е истина)	Няма грешка	Грешка „Тип 1”
Последователността е не-случайна ( $H_1$ е истина)	Грешка „Тип 2”	Няма грешка

В реалните тестове се избира ниво на значимост  $\alpha$ . В този случай нулевата хипотеза  $H_0$  се приема, ако *P-стойност*  $\geq \alpha$ . Ако *P-стойност*  $< \alpha$ , то  $H_0$  се отхвърля, което показва, че последователността е неслучайна. Така дефинираният параметър  $\alpha$  показва вероятността за „грешка тип 1”. Обикновено  $\alpha$  се избира в интервала [0.001, 0.01]. В извършените тестове е избрана стойност  $\alpha = 0.01$ . Тази стойност показва, че е допустимо една от всеки 100 изследвани последователности да се отхвърли. Изпълнението на неравенството *P-стойност*  $\geq 0.01$  означава, че последователността ще се счита за случайна с доверителна вероятност 99 %. Обратно, ако *P-стойност*  $< 0.01$ , то се прави заключение, че последователността е неслучайна с доверителна вероятност 99 %. В този случай са възможни два типа грешки (Таблица 1 [9]). Заключение, грешка „Тип 1”, се получава в случаите, когато последователността в действителност е случайна, а тестът отхвърля нулевата хипотеза  $H_0$ . Грешката „Тип 2” възниква в случаите, когато изследваните данни са неслучайни, а тестът приема  $H_0$ .

За определяне на случайността на произволно дълга двоична последователност, генерирана от *pGSSG* генератора, е използван пакетът „*NIST Test Suite statistical package*” [6]. Той съдържа 15 функционални теста, като някои от тях се състоят от множество подтестове. Основната цел на всеки от тестовете и характеристиките на последователността, които те откриват, накратко са систематизирани по-долу. Повече информация за математическите основи на тестовете може да се намери в [8, 9]. Петнадесетте теста са: Честотен тест (Монобит), честотен тест в блок, тест с кумулативни суми, сериен (последователен) тест, тест за най-дългата серия от единици в блок, тест за определяне на ранга на двоична матрица, тест с дискретно преобразуване на Фурие (спектрален тест), тест за съвпадение на не-припокриващи се шаблони, тест за съвпадение на припокриващи се шаблони, универсален статистически тест на Мюрер, приблизителен ентропиен тест, тест на средните отклонения, вариантен тест на средните отклонения, сериен тест и тест за линейна сложност.

**Diehard** е първият придобил популярност и широка употреба тестови пакет. Разработен е от George Marsaglia [5], изследовател в областта на генерирането на псевдослучайни числа и тестовете за случайност. Пакетът се състои от над 10 теста, повечето от които са основани на негови изследвания, изчисляващи статистически стойността на случайността, като открива често срещаните недостатъци в изследваните генератори. Въпреки това, след публикуването му са открити множество пропуски като липсата на тестове на ниво битове и тестове, свързани с криптографската сила. Друг недостатък е липсата на възможност на параметризиране на тестовете и съответно по-малката гъвкавост при провеждане на тестовете.

През 2004 е публикуван **DieHarder**, разработен от Robert Brown [1], който имплементира обобщени версии на всички Diehard тестове, като ги разширява с някои от статистическите тестове на NIST и разработени от автора допълнителни тестове. Инструментът е изцяло базиран на GSL интерфейса за генератори на случайни числа. Една от основните цели на DieHarder е да направи възможни тестовете над много дълги последователности, затова и тестовите данни се съхраняват като файлове със случайни числа, като при Diehard броят им е ограничен до едва 10 милиона.

Подобренията в DieHarder могат да се обобщят в три групи. Първата е свързана с теста на Коломогоров-Смирнов, при който много генератори преминават успешно в Diehard, но всъщност тяхното разпределение на *p*-стойностите не е еднородно. Второто подобрение се състои в промяната на броя семпли в един статистически тест за single-run. Множество от тестовете на Diehard изследват припокриването на последователности от битове. Припокриващите се последователности не са независими и това трябва да се вземе предвид при определяне на стойността на автокорелацията при различно отместени семпли. Третото подобрение се изразява във възможността за използване на тестови файлове, в които да се запамятат и тестват милиони стойности. Тестовият пакет продължава да се поддържа и да се допълва постоянно, като последна версия е Version 3.31.1 от 2018.

Друг набор от тестове за случайни числа е **ENT**. Най-отличаващият му тест е пресмятане на числото  $\pi$  по метода на Монте Карло. Освен него съдържа още ентропиен тест, Chi-square тест, тест на средно аритметичното и коефициент на серийна корелация. Тестовата програма има слаба документация и малко параметри за настройка на тестовете. Подобно на Diehard работи с файлове, но не разполага с вградени рандом генератори [11]. Пакетът не е обновяван от 2008 година и е по-скоро приложим за обучителни цели отколкото за изследователски такива.

Още един софтуерен пакет на ANSI C за изследване на генератори на случайни числа е софтуерната библиотека **TestU01**. Тя е дело на Pierre L'Ecuyer и Richard Simard [4], като в основата си се базира на тестовете, описани в Knuth [3], и в следствие се разраства до библиотека, имплементираща множество подобрени и авторски тестове. С някои изключения тя напълно припокрива тестовете в Diehard и статистическите тестове на NIST. За разлика от

ENT притежава богата документация, разглеждаща в детайли отделните тестове и теорията зад тях. Тестовите са разделени в модули, като някои от тях са *smultin* (клас тестове, базиран на мултиноминално разпределение), *sentrop* (клас ентропийни тестове), *snprair* (тестове, базирани на разстояние между най-близки точки), *sknuth* (имплементация на тестовите от Кнут), *smarsa* (тестове от Diehard), *svargia* (набор от тестове, базирани на проста статистика) и други. Освен тестовите библиотеката имплементира широк набор от известни генератори като линейен конгруентен генератор, генератори използвани в криптографията като AES OFB, Mersenne Twister, използван при Монте Карло симулациите [4]. Последната версия на библиотеката е документирана през месец май 2013 г.

**Таблица 1.** Софтуерни инструменти за тестване

Инструмент	Последна версия	Дата на последно обновяване
<i>NIST STS</i>	SP 800-22rev1a Sts-2.1.2	Април 2010 (спецификация) Юли 2014 (софтуер)
<i>Diehard</i>	Не е достъпен	1995
<i>DieHarder</i>	V 3.31.1	2018 (спецификация и документация) Октомври 2011 (софтуер)
<i>ENT</i>		Януари 2008
<i>TestU01</i>	TestU01-1.2.3	Август 2009

Данните за разгледаните тестови пакети са обобщени в таблица 1. След обстойното проучване може да се направи извод, че при изследването на генератори на псевдослучайни числа за пълнота е добре да се използва повече от един тестови инструмент. NIST статистическите тестове са най-актуални като софтуер, докато DieHarder и TestU01 притежават по-широк набор от тестове. Съответно популярните преди години Diehard и ENT могат да се считат за остарели и могат да бъдат пропуснати в методологията за оценка на случайността.

### 3. Оценка на недвоични последователности

Нуждата от създаването на нови генератори на псевдослучайни числа не престава да бъде все така актуална. С навлизането на интернет на нещата става необходимо оптимизирането им за използване във вградени системи и устройства с ниска изчислителна мощност. Това налага конструирането на нови генератори за нуждите на т. нар. лека криптография (*lightweight cryptography*) и съответно тяхното изследване. Тук ще предложим методология за оценка на недвоичните псевдослучайни последователности, получени от подобни генератори.

#### *Начин на тестване*

Можем да разграничим два начина на тестване на генераторите в зависимост от режима на тяхната работа, а именно онлайн и офлайн. Онлайн тестването се провежда в реални време, докато генераторът работи и изходните числа се тестват едно по едно. При офлайн метода изходните данни се съхраняват като последователност от числа, за да се анализират по-късно. В практиката са се наложили офлайн тестовите, тъй като те позволяват разглеждането на последователността като цяло или на отделни аспекти от нея. Важно условие е изследваната последователност да е с достатъчно голяма дължина, за да може при успешно преминаване на тестовите за случайност да се заключи, че и останалите изходни последователности имат същите характеристики.

Предложената по-долу методология предвижда офлайн тестване с цел подлагането на изходните последователности на многократни тестове за оценка.

### Методология за оценка

По-долу е представена обобщената методология за оценка на случайността на недвоични псевдослучайни последователности. От избраните тестови пакети единствено DieHarder поддържа работа с цели числа и това налага преобразуването им в двоичен код за целите на другите два тестови пакета.

Таблица 2. Формат на входните данни

Инструмент	Форматиране на входните файлове	Числа
<b>NIST STS</b>	Двоични последователности записани като <ul style="list-style-type: none"><li>текстови, ASCII символите '0' или '1' или</li><li>двоични данни, където всеки байт съдържа 8 бита (нули и единици)</li></ul>	двоични
<b>DieHarder</b>	Три вариант на вход: <ul style="list-style-type: none"><li>raw binary stdin – поточен интерфейс за получаване на данните от стандартния вход в двоичен вид</li><li>raw binary file – двоичен файл</li><li>file input ascii – текстови файл, съдържащ случайните числа обикновено от тип uint със ascii форматиране. Файлът съдържа специфичен хедър, а числата се записват по едно на ред.</li></ul> Поддържа файлове с големина над 2GB	двоични или недвоични
<b>TestU01</b>	Приема само 32-битови входни данни (числа с плаваща запетая) и ги интерпретира като стойности в интервала [0, 1). Няма специфично форматиране, числата се разделят с кой да е whitespace символ. Файлът трябва да съдържа само валидни реални числа. * <i>Представянето на числата като дробни въвежда възможност за грешка от закръгление и загуба на информация при най-младшия бит.</i>	с плаваща запетая

#### Стъпки:

##### 1. Избор на генератор за тестване.

Необходимо е да се избере конкретен генератор на недвоични псевдослучайни последователности, който ще бъде обект на оценка. Някои тестове бяха проведени с r-ичен Обобщен самосвиващ генератор.

##### 2. Определяне на набор от стойности за параметрите на генератора, ключове и начално състояние за конкретния тест.

За да бъде валидно изследването на генератора, трябва да бъдат тествани широк набор от изходни последователности. В зависимост от спецификата на генератора могат да се изберат различни начални състояния, настройки, свързани с дължината на периода на изходната последователност, и други. Създава се набор от настройки.

##### 3. Стартиране на генератора и запис на изходната последователност от псевдослучайни числа.

Въвежда се всяка предвидена настройка на генератора и резултатът от неговата работа се записва във файл. Практически тестове показват, че е препоръчително да се генерират милиони числа (над 30 млн.) или над 120MB двоични данни, за да се постигне автентичност на резултатите.

##### 4. Избор на софтуерен тестови инструмент.

Избира се един от описаните в секция 2 софтуерни тестови инструменти, а именно статистическите тестове на NIST, DieHarder и TestU01.

5. *Подготовка на файл в зависимост от изискванията за форматиране на входните данни*

В таблица 2 са описани спецификите на входните файлове на трите тестови инструмента. В съответствие с тях е възможно да се наложи преобразуване на получените от изследвания генератор последователности. Така например, за получаването на двоична последователност всяка  $p$ -ична цифра  $i$  се преобразува до двоичен код (1) [10].

$$(i-1) + \frac{2^{\lceil \log_2 p-1 \rceil} - p - 1}{2} \quad (1)$$

За целите на TestU01 всяка  $p$ -ична цифра  $i$  се преобразува до дробно число в интервала  $[0, 1)$  посредством (2)

$$\frac{i}{p} \quad (2)$$

6. *Провеждане на тестовете със софтуерния инструмент.*

Стартиране на софтуерния инструмент съгласно неговите инструкции за работа, необходими настройки и подавайки създадения в стъпка 5 входен файл.

7. *Преглед и оценка на резултатите*

- При наличие на неуспех се прекратява тестването и генераторът се отхвърля.
- При успешно преминаване на тестовете от даден инструмент се преминава към стъпка 4 за следващия инструмент.
- При успешно преминаване на всички тестове в тестовите пакети генераторът се счита за такъв с достатъчно високо ниво на случайност.

8. *Опционална стъпка: Провеждане на допълнителни тестове.*

Тъй като емпиричното изследване за случайност не може да даде категоричен отговор и реално пълната случайност не е постижима, е препоръчително да се разгледат и други тестове, предложени през последните години от научните общности.

**Таблица 3.** Двоично представяне на  $p$ -ичната цифра [10]

$p$ -ична цифра	$p = 3$	$p = 5$	$p = 7$	$p = 11$	$p = 13$
1	0	00	001	0011	0010
2	1	01	010	0100	0011
3	-	10	011	0101	0100
4	-	11	100	0110	0101
5	-	-	101	0111	0110
6	-	-	110	1000	0111
7	-	-	-	1001	1000
8	-	-	-	1010	1001
9	-	-	-	1011	1010
10	-	-	-	1100	1011
11	-	-	-	-	1100
12	-	-	-	-	1101

**Допълнителни тестове**

Както вече беше уточнено, природата на псевдослучайните числа не позволява те да бъдат определени като напълно случайни, а следва да се избере ниво на случайност, на което те трябва да отговарят. Разгледаните по-горе тестови инструменти са утвърдени и към настоящия момент се считат за достатъчни, за да се направи оценка дали един генератор

отговаря на изискванията за случайност. Изследвания през последните 5 години [2, 7, 12, 13] предлагат нови тестове, базирани на редица аритметични закони.

В свое изследване Уанг и Никол [12, 13] показват, че е лесно да се конструират генератори, чиито изходни последователности очевидно не отговарят на изискванията за равномерно разпределение, но в същото време успешно преминават статистическите тестове на NIST. Те предлагат прилагането на допълнителни тестове, базирани се на две основни ограничаващи теореми за двоични низове – централна ограничаваща теорема и закон на повторения логаритъм (LIL).

През 2017 Павел Лорек [7] публикува нов статистически тест, базиран на аркус синус закона за „random walk“. Подобно на LIL тестовете, изходната последователност се разглежда като двоичен низ вместо последователност от числа. Методът определя качеството на генераторите чрез измерване на статистическо разстояние между емпиричното разпределение на разглежданата характеристика за генерираната изходна псевдослучайна последователност и нейното теоретично разпределение за истински случайни двоични последователности. Авторите демонстрират намирането на недостатъци в някои генератори, които преминават успешно LIL тестовете.

Отново през 2017 друг изследовател, Хироши Харамото, предлага експериментален метод за откриване на дефекти в статистическите тестове, използвайки тристепенен тест [2]. Така въпросът за изчерпателното тестване за случайност на генерираните последователности остава отворен.

#### **4. Заключение**

Необходимостта от изследване на случайността на последователностите, получени от различни генератори, се обуславя от тяхното широко приложение в редица области като криптографията, симулацията, електронните и хазартни игри. Теорията около тези последователности се развива основно през последния половин век, като огромна част от изследванията са посветени на двоичните последователности. От разглежданите спецификации на тестове и създадените софтуерни инструменти за прилагането им можем да направим извод, че те откриват голям процент от отклоненията от равномерното разпределение на последователностите. Препоръчително е ползването на инструментите NIST STS, Dieharder и TestU01 в комбинация.

Авторите предлагат методология за оценка на недвоични последователности чрез офлайн тестване. В нея е отразена необходимостта от преобразувания на недвоичните числа, за да се съвместят със спецификата на входните данни на някои от тестовите инструменти.

Не на последно място остава отворен въпросът за имплементирането и провеждането на допълнителни тестове за оценка, които все още не са се наложили като част от утвърдените софтуерни инструменти.

#### **Благодарности**

Тази работа е резултат от проект, финансиран от Фонд Научни Изследвания към Министерството на образованието и науката на България, по „КОНКУРС ЗА ФИНАНСИРАНЕ НА НАУЧНИ ИЗСЛЕДВАНИЯ НА МЛАДИ УЧЕНИ – 2016 г.“ [Номер на договора ДМ07/5 – 15.12.2016]

#### **Литература**

[1]. Brown, Robert G., Eddelbuettel, D., Bauer, D., Dieharder: A Random Number Test Suite, Version 3.31.1, 2018, <https://webhome.phy.duke.edu/~rgb/General/dieharder.php>

- [2]. Haramoto H., Checking the Soundness of Statistical Tests for Random Number Generators by Using a Three-Level Test. arXiv:1710.06033, 2017.
- [3]. Knuth, Donald E., The Art of Computer Programming, Addison Wesley Longman, Inc., 3 edition, 1997
- [4]. L'Ecuyer, P., Simard, R., TestU01: A Software Library in ANSI C for Empirical Testing of Random Number Generators, 2013, <http://simul.iro.umontreal.ca/testu01/guideshorttestu01.pdf>
- [5]. Marsaglia, G., The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness, 1995, <http://www.stat.fsu.edu/pub/diehard/>
- [6]. NIST. Test suite. 2010. <http://csrc.nist.gov/groups/ST/toolkit/rng/>
- [7]. Paweł Lorek, Grzegorz Łoś, Filip Zagórski, and Karol Gotfryd. On testing pseudo random generators via statistical tests based on arcsine law. [http://www.math.uni.wroc.pl/~lorek/papers/main\\_arcsine\\_AMC\\_SUBMITTED.pdf](http://www.math.uni.wroc.pl/~lorek/papers/main_arcsine_AMC_SUBMITTED.pdf), 2017.
- [8]. Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E., Leigh, S., Levenson, M., Vangel, M., Banks, D., Heckert, A., Dray, J., and Vo, S., NIST Special Publication 800-22rev1a: A Statistical Test Suite for the Validation of Random Number Generators and Pseudo Random Number Generators for Cryptographic Applications, April 2010, <http://csrc.nist.gov/groups/ST/toolkit/rng/documents/SP800-22rev1a.pdf>.
- [9]. Soto J., Statistical Testing of Random Number Generators, National Institute of Standards & Technology, Proceedings of the 22nd National Information Systems Security Conference, October 1999, p. 12.
- [10]. Tasheva, A., Nakov, O., Trifonov, T., „Statistical Analysis of the Sequences Produced by the p-ary Generalized Self-Shrinking Generator“, Доклад на международна научна конференция 17th WSEAS International Conference on COMPUTERS (part of CSCC '13), Rhodes Island, Greece, 16-19 юли 2013.
- [11]. Walker, John. ENT, Fourmilab, 2008, <http://www.fourmilab.ch/random/>
- [12]. Wang, Y., On the Design of LIL Tests for (Pseudo) Random Generators and Some Experimental Results, arXiv:1401.3307v1, 2014
- [13]. Wang Y., Nicol T., Statistical Properties of Pseudo Random Sequences and Experiments with PHP and Debian OpenSSL, Computer Security - ESORICS 2014. Lecture Notes in Computer Science, vol 8712. Springer, 2014

**За контакти:**

доц. д-р Антония Т. Ташева  
кафедра „Компютърни системи”  
Технически университет - София  
E-mail: atasheva@tu-sofia.bg

# A SYSTEM FOR WEBSITE PUBLISHING IN A WEBSITE BUILDING SYSTEM

Dimitar Z. Dimitrov

**Abstract:** The paper describes a conceptual model of an architecture for a system for website publishing, used by a website building system for creating and publishing websites online, based on website data structures, created using a visual website editor. The paper describes a software architecture divided into several subsystems and processes aiming to transform the website's data to a ready to use website and making it available online through a set of subsystems. There is no other published research of architecture of this kind.

**Keywords:** website builder system, online publishing, website editing, website delivery, website deployment, website data management, website versioning, multiwebsite systems, multitenant systems

## Introduction

Interactive website building systems are now commonly used to design dynamic websites. They are usually complex systems, implementing a set of architectures and subsystems. Modern online website building systems usually implement a multitenant architecture [1], supporting multiple users, and in some systems a single user can support multiple website, which can also be described as “multiwebsite system” [2].

This paper proposes a solution for an architecture for a subsystem, designed to manage the process of website online publishing in a website building system. The system includes a website caching system, website dependency management system, website deployment system and in addition makes use of other systems in a website building system environment like an image optimization system, a website data management system and a routing system.

The website publishing system takes the saved website data from the website visual editor, transforms it to a corresponding code (HTML, CSS, JS, etc.), includes all required libraries, depending on the used components, optimizes all images, and makes it available online through a website delivery system. The website delivery system in term includes a routing system and a content delivery network (CDN).

A website building system can grow to hundreds of millions of websites, all working simultaneously. Each of the websites will on their term be accessed thousands of times. The website publishing system will be using a website data management system to be able to handle the big data [3].

The paper proposes a conceptual model of an architecture for a website publishing system in a website building system, which at this moment is the only publicly proposed architecture of its kind. The architecture is implemented in a website building system, operating for more than 2 years as part of a website building online platform of an international public listed company.

## Website publishing system in a website building system

### 1.1. Field of research

There is very little disclosed information on website building systems and architectures. Most of which is disclosed through patents and few of which is disclosed openly, usually not sensitive information with relatively small significance. The present paper relates to website building systems



generally and to editing tools in particular. Websites created using a website building platform are typically created and edited by users and then accessed by end users.

A website building system typically allows the user to create and edit applications. A visually designed application consists of pages, containers, and components. Pages are typically displayed separately and contain components. Components are typically arranged in a hierarchy of containers (single page and multipage) inside the page including atomic components. A multipage container may display multiple minipages [4].

Pages may use templates either general page templates or component templates. Specific cases for templates include the use of an application master page containing components replicated in all other regular pages, and the use of application header/footer (which repeat on all pages).

Components typically have attributes (e.g. size, position, color, frame type, etc.) and possibly content (e.g. contained text in a text component). Components may be fixed (such as a given geometrical shape) or based on a content management system (CMS) (such as a component containing a picture selected from a database of available pictures). Such a database may be extended by the user through importing (in this example) additional pictures. Components may also be based on user-specified content, (such as a text area into which the user enters text) or be based on external information (static or dynamic), such as a RSS feed displaying information from an external data source or the content of a given internet page. Component appearance can be modified by resizing, rotation and similar operations. Their appearance and behavior can also be modified by various modifiers, such as blurring and animation. A specific set of attributes for a given components (e.g. a button having a given shape, color, text font) is referred to as component configuration. It will be appreciated that the component configuration definition does not include its content (e.g. the buttons' text label), position and size [5].

## **1.2. Related Works**

Most of the research in this niche field of website builder systems is conducted by corporations and is not disclosed as it is sensitive Intellectual Property (IP). Some research is disclosed as approved published patents and patent applications.

Revealed patents in the website builder system industry do not reveal or propose an architecture for a website publishing system. There are some good researches in the fields of some of the subsystems in place.

Towards a feature for the caching system a research for HTML automatic table layout [6] has been conducted by K. Marriot and P. Moulder. It provides a solution for one of the possible caching approaches for adaptable design to the user's viewport. The research suggests a heuristic algorithm, using a greedy strategy that starts from the widest reasonable layout and repeatedly chooses to narrow the column for which narrowing leads to the least increase in table height. Although table layout is an option, it is a quite limiting layout.

The proposed system can use either responsive design, or absolute positioning. The experiments conducted are implemented on a system with absolute positioning and mobile adaptive design. And it does not include tables for layout.

In the field of multimedia (images, videos, etc.) publishing there has been plenty research and multiple services available, that can be integrated into the suggested system. An interesting research for a flexible environment for multimedia management and publishing has been conducted [7]. It describes a system called IM3I, which can be defined as an intelligent and adaptable multimedia repository software.

Widely-used services for an online file publishing and storage API are cloud-based services like the Amazon S3 service, part of the Amazon Web Services (AWS) platform [8]. The proposed system has been implemented using AWS S3 service for online file storage, but it is abstracted so it can be implemented on any similar type of storage service.

Moving forward to research on website data structures, there are a few researches all going in the same direction, and our website data management system is also implementing a similar, more sophisticated model. A separation between content and presentation [9, 10] is suggested, as in our case we also separate the settings of each component. Also, other researchers interpret the presentation as a layout of the whole webpage or a code fragment. In our system, the presentation is separated per component, and is stored in the data structure of the website along the content and the settings, and the website data structure is used by the caching system as one whole entity.

As a publishing system, available to the general crowd, using user generated content, measures for securing intellectual property rights must be taken, when considering such a system. A research on intellectual property aspects of web publishing has been described by H. Kienle and D. German [11]. Automated systems may be created for unpublishing automatically any content, which has been detected to have breached copyrights. On the other hand, the same applies for abusive content. Systems using artificial intelligence are now available to detect abusive content and can automate this action. Another topic is the possibility of censorship, which must be avoided, and there is also research of such frameworks [12].

Perhaps the most relative research to the currently suggested system is suggesting a fragment-based approach for efficiently creating dynamic web content, conducted by an IBM research team [14]. The approach they suggest is focusing more to the caching system, by offering approaches to cache fragments, define relations by object dependency graphs. The described approach brings some clarity to the template caching problem. Similar approach has been established by major web development frameworks like the PHP framework Symfony. Which in term confirms the right direction of the caching approach by fragmenting the sources.

Another research field is when we go down to the presentation layer of each component. A component can be represented as an application and can have its own data structure to be as an executable code. A similar approach is suggested, although instead of components they research creating interactive data-driven web applications [15]. In this research, they go further in describing a dynamic part of an application being the data storage. In the system, suggested in this paper, the dynamic content has its own system and is managed separately, while the publishing system organizes its workflow in the caching and delivery systems.

Another research by the MIT Computer Science and Artificial Intelligence Laboratory shows a data publishing approach based on an exact description of a data structure [16]. This is a very early approach to web publishing based on data structures. While in a modern website building system a visual approach to creating website data structures, this research is based on coding configurations, which are being interpreted in the process of publishing. The layers of presentation and configuration is not available and may be possible through coding.

A simpler approach on publishing using data structures has been suggested in a research for a niche system for digital books, magazines and newspapers publishing [17]. It again uses data structures written by users, but in this case, it does not have a separate presentation layer at all but has one presentation for all users. It generally suggests a caching system for the generated layout from the data structures of the authors. Interesting part of the paper is utilizing the capabilities of browser caching, which is not part of the caching system, but is a small piece, included in the website delivery system.

### **1.3. Requirements and Dependencies**

The website building systems can be implemented to execute in an online, offline or partially online environment. The proposed system is most relevant to online implementations, using multitenant and multiwebsite architecture, with a SaaS (Software as a Service) [18] delivery method. The website publishing system requires a minimum set of subsystems of a website building system.

### 1.3.1. Website data management system [19]

A system divided into a Data Access Layer (DAL) which abstracts the direct connection and querying the selected database, and a data object mapping layer like Object Relational Mapping (ORM) [20] and Object Document Mapping (ODM) [21] layer. A good example of an open source framework for the required abstraction layer is the Doctrine project [22], which were also used in the experiments.

The data management is a part of a separate system in the website building system, but a requirement for high performance of the data management system is mandatory for the DVM system to operate, being designed to work in real time.

NoSQL [23] databases are key to supporting Big Data applications, since they enable handling large quantities (i.e. volume) of highly-variable (i.e. variety), user-generated contents while guaranteeing fault tolerance, availability (i.e. velocity) and scalability. [24]

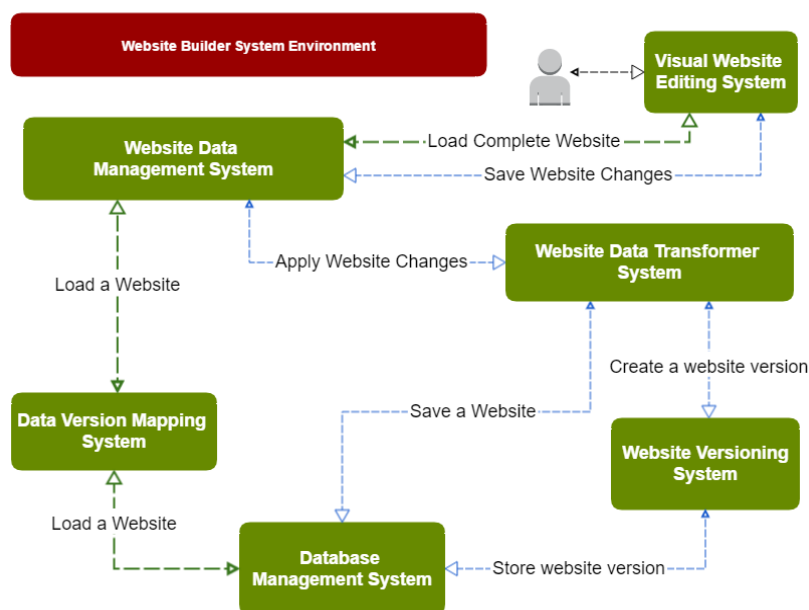
### 1.3.2. Dependency manager system

This is the system that manages all the requirements in a website builder system, based on what the different subsystems require. In this case, this system will make sure that all the currently described systems are included and operational before executing the WDMS.

### 1.3.3. Website versioning system

This system is used for backing up website data and can be a part of the currently described website data management system or can be a separated system. The website versioning system gets as input a data object with a website interface and makes a copy of it, usually stored in a separated archive database.

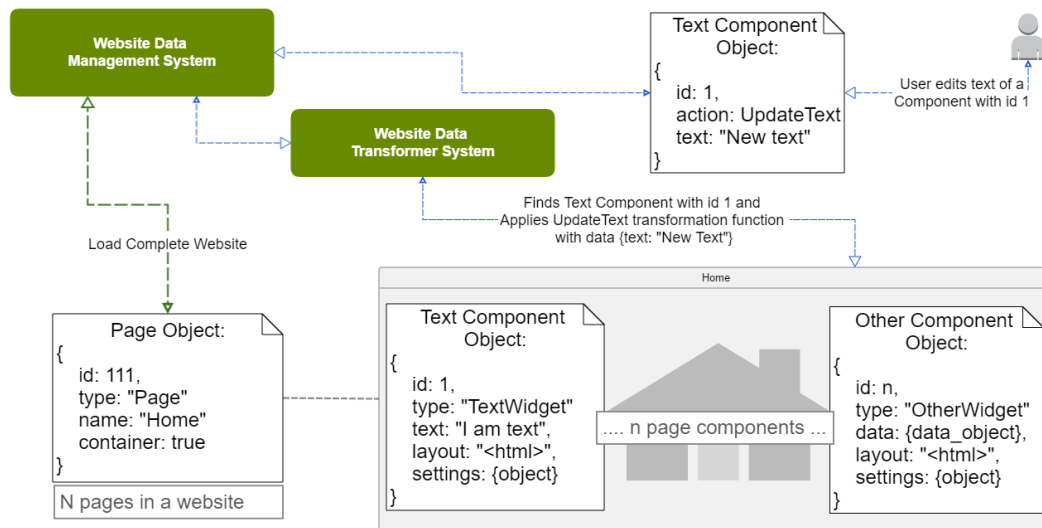
Usually website versioning is called on each change of a website and is used as website history. A functionality for website restoring is introduced to users of some website builder systems. It generally reverts the current state of the website data to the state it was archived at a certain moment.



**Fig 1.** Architecture and workflow of systems for website data management, data version mapping, database management, website versioning and website data transformation

Described in Figure 1 is the workflow of the two main operations in a website data management system. The operations for loading a website into the visual editor, and then the operation for saving the edited website data in the website builder system.

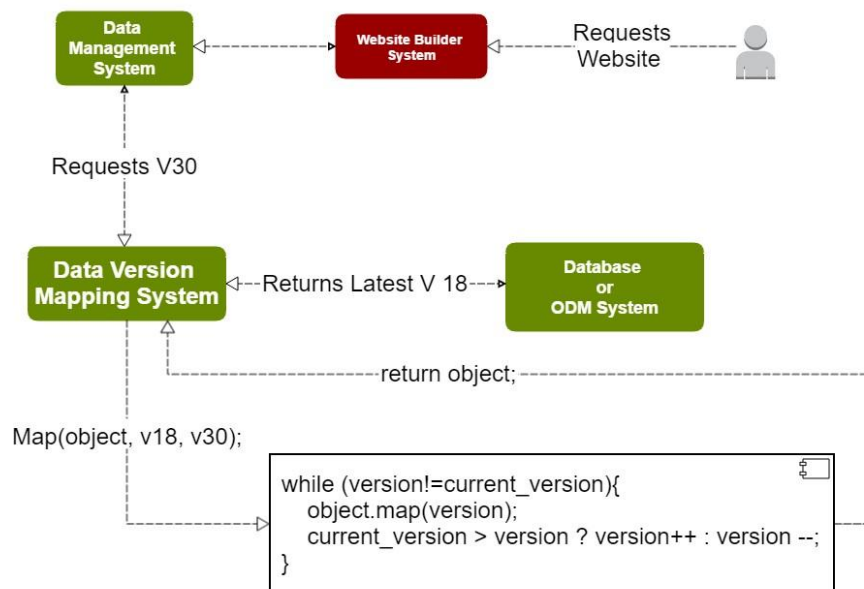
### 1.3.4. Website Data Transformer System



**Fig. 2.** An example of how the data moves through the workflow. In more detail, it visualizes the workflow of a website transformer system

### 1.3.5. Data version mapping system (DVM) [25]

This system utilizes methods for “lazy data migration” [26], applicable to big data applications, required to operate in real time.

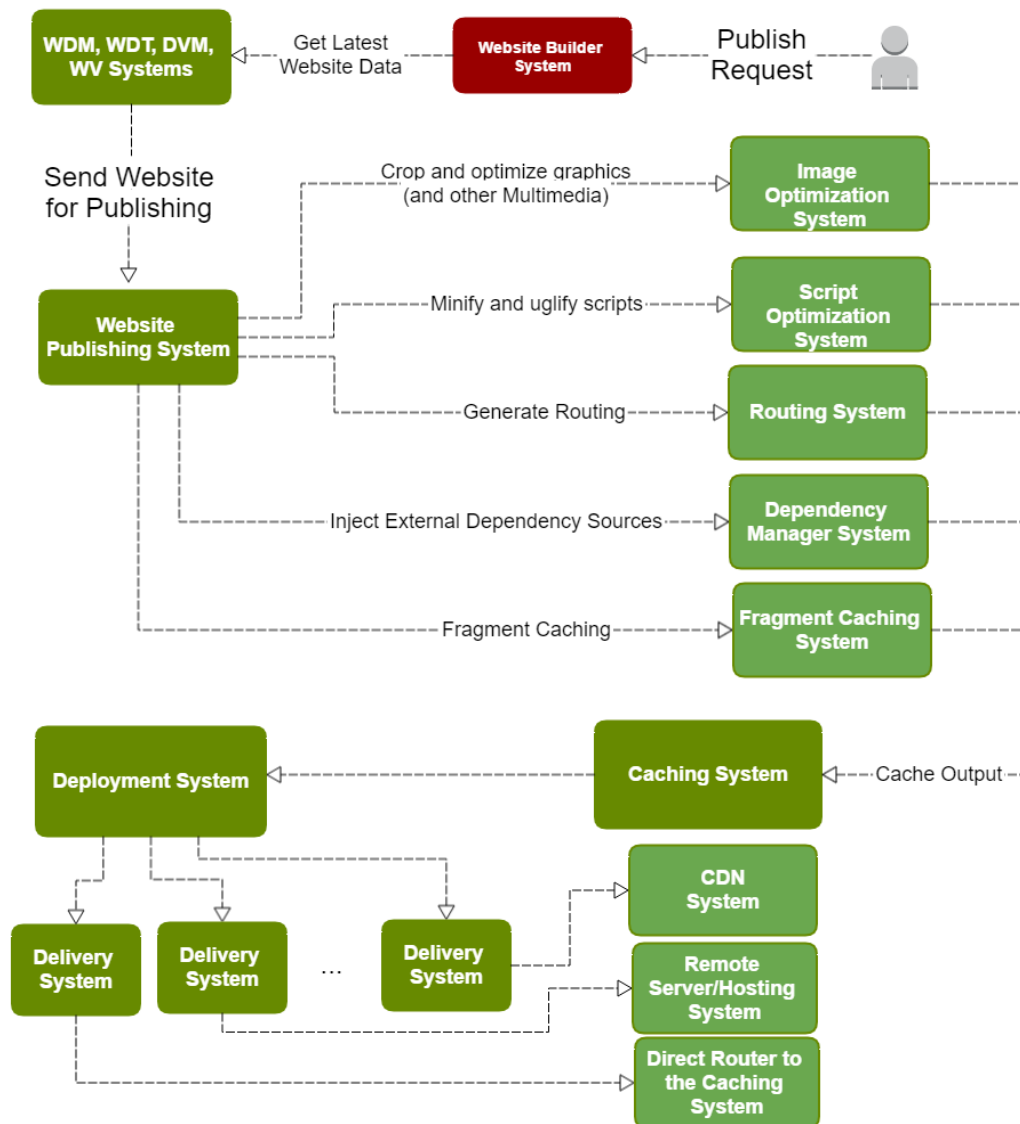


**Fig. 3.** General representation of how the Data Management System interacts with an example DVM system

A DVM system enables the existence of many versions of a component’s data, enabling a website system to be able to run different versions of a system simultaneously. To be able to release new versions rapidly or to revert to older versions seamlessly. Also, possible A/B testing of

different release branches of the same component or feature, while keeping the data available for all the versions of the system.

## WEBSITE PUBLISHING SYSTEM ARCHITECTURE



**Fig 4.** Architecture and workflow of a website publishing system

Figure 4 describes system requests as an input an instance of a website data structure, provided by the Website Data Management (WDM), Website Data Transformation (WDT), Dynamic Version Mapping and Website Versioning Systems. The publishing system orchestrates the publishing process by using a set of subsystems in parallel. Starting from the Multimedia (Image/Video, etc.) optimization system, which crops the images to the exact sizes, optimizes all media for web, then a Script Optimization System is managing all generated scripts to be of optimal minified size, and go through a process known as uglification, which in general terms parses scripts and replaces references, e.g. variable or class names with short symbols. This helps for reducing the script sizes dramatically and helps protect and secure it, by decreasing the readability of the code to minimum.

One of the most important systems is the Dependency Manager System. It keeps track of all dependencies for the objects being published and injects them. It can be used fonts as the simplest

example, which the DM system makes sure they are injected as files and as source code in the scripts of the published pages. For that reason, the website builder system must use a DM system to keep track of all object dependencies. Another example is a simple link dependency. You can have a link from one page to another, but if you delete the linked page, the link will not work. Here the dependency manager should give an alert for that case, while the person is trying to delete a dependency page.

The Fragment Caching System (or Template Caching System) is the biggest system of the set, but is not a subject to this research, as discussed in the section related work, there has been many researches done in this direction, and we are using one of the many proposed caching systems. A fragment caching system is now widely used in major development frameworks. In a website builder system, a fragment can be a component or a whole page itself. When a user updates just one component, it can update the cache to its cached fragment, and only update the page fragment it is part of again by using a dependency manager, managing an object dependency graph.

An interesting subset of a fragment caching system is caching dynamic fragments. E.g. when you have dynamic contents, like an online shop, adding new products. When we want to cache this content fragment, there must be a system in place to trace and invalidate this cache based on the object dependency graph, after each update on each dynamic object. E.g. changing the price of a product should invalidate the product's details page, and all listing fragments, cached, which it is part of. All subsystems have an output, which is cached by a caching system. The caching system can be file storage, database, in-memory storage, or any specialized caching system like Memcache or Redis.

The Deployment System has the task to deploy the cached website to a Delivery System or a set of Delivery Systems if required by system architecture. The Delivery Systems delivers the website to the end users. When the cached website is deployed to a delivery system, we can say that the publishing is completed. The delivery system can have a caching system of its own. In the examples in Figure 4, one of the delivery systems includes a Content Delivery Network (CDN) to cache and deliver the website. Another delivery system is a remote webserver, to which the deployment system has deployed the cached website by ftp or other method. A third option is that the delivery system can be just a relay proxy directly to the internal caching system of the website builder, which is the simplest solution. In this case, the deployment system can be excluded or can be part of the caching system itself.

## Conclusions and future work

Major website builder systems today have tens and hundreds of millions of websites per system. And the niche is growing rapidly. Research in the field is being done corporately and very few is made public, usually by patent applications. The research found in the field is in most cases not older than a couple of years, which makes continuing the research in the field of website building applications even more interesting.

## References

- [1]. M. Kapuruge. Service Orchestration as Organization: Building Multi-Tenant Service Applications in the Cloud - Morgan Kaufmann, 2014
- [2]. Dimitrov D. Standardization based on multiwebsite system integration – International Science and Technology Conference, Dubai, 2012
- [3]. V. Mayer-Schönberger. Big Data: A Revolution That Will Transform How We Live, Work, and Think - Eamon Dolan/Mariner Books, 2014

- [4]. R. Ben-Aharon, N. Abrahami. System and method for the creation and use of visually-diverse high-quality dynamic visual data structures, US 2016/0357527 A1 - United States Patent Application Publication, 2016
- [5]. A. Geva, O. Gartzman, G. Kaplan. System and method for the generation of an adaptive user interface in a website building system, US 2016/0357527 A1 - United States Patent Application Publication, 2016
- [6]. K. Marriot, P. Moulder. HTML Automatic Table Layout - ACM Digital Library, 2013
- [7]. M. Bertini, G. Ioannidis and E. Bijk. A flexible environment for multimedia management and publishing - ACM Digital Library, 2017
- [8]. Y. Wadia. AWS Administration - The Definitive Guide - Packt Publishing, 2016
- [9]. D. Dudek, H. Wiecek. A Simple Web Content Management Tool as the Solution to a Web Site Redesign - ACM Digital Library, 2003
- [10]. Adi Omari, S. Shoham. Lossless Separation of Web Pages into Layout Code and Data - ACM Digital Library, 2016
- [11]. H. Kienle, D. German. Intellectual Property Aspects of Web Publishing – ACM DL, 2004
- [12]. M. Waldman, L. Cranor. The Architecture of Robust Publishing Systems - ACM DL, 2001
- [13]. J. Simonson, D. Berleant. Content Permanence Via Versioning and Fingerprinting - ACM Digital Library, 2000
- [14]. J. Challenger, P. Dantzig, A. Iyengar and K. Witting. A Fragment-Based Approach for Efficiently Creating Dynamic Web Content - ACM Digital Library, 2005
- [15]. L. Verou, A. Zhang, D. Karger. Mavo: Creating Interactive Data-Driven Web Applications by Authoring HTML- ACM Digital Library, 2016
- [16]. D. Huynh, D. Karger and R. Miller. Exhibit: Lightweight Structured Data Publishing - ACM Digital Library, 2007
- [17]. R. Aamulehto, M. Kuhna and J. Tarvainen. Stage Framework - An HTML5 and CSS3 Framework for Digital Publishing - ACM Digital Library, 2013
- [18]. A. Fox, D. Patterson. Engineering Software as a Service: An Agile Approach Using Cloud Computing - Strawberry Canyon, 2013
- [19]. Dimitar Dimitrov: System for website data management in a website building system, CompSysTech'2017 Proceedings, ACM ICPS, vol.1396, doi>10.1145/3134302.3134313, p.p. 211-218, 2017
- [20]. A. Kurniawan. Object-Relational Mapping (ORM): MongoDB, Mongoosejs and Node.js - PE Press, 2012
- [21]. Doctrine Project Organization. Blending the ORM and MongoDB ODM - <http://docs.doctrine-project.org/projects/doctrine-mongodb-odm/en/latest/cookbook/blending-orm-and-mongodb-odm.html>, 2014
- [22]. Doctrine Project Organization. <http://www.doctrine-project.org/>
- [23]. M. Scavuzzo, D. Tamburri. Providing Big Data Applications with Fault-Tolerant Data Migration Across Heterogeneous NoSQL Databases, 2016
- [24]. S. Scherzinger, M. Klettke, Managing Schema Evolution in NoSQL Data Stores - DBPL'13, 2013
- [25]. Dimitar Dimitrov, Elena Racheva: System and method for dynamic data version mapping in a website building system, Computer Science and Technologies Journal, Technical University – Varna, №1/2017, pp.77-82, 2017
- [26]. S. Scherzinger, U. Störl, M. Klettke. A Datalog-Based Protocol for Lazy Data Migration in Agile NoSQL Application Development – ACM Digital Library, 2015

**For contacts:**

PhD candidate Dimitar Z. Dimitrov  
 Technical University of Varna, Bulgaria  
 E-mail: dimitrov@oxy.com

# A SYSTEM FOR OPTIMIZATION OF IMAGES IN WEBSITE BUILDING SYSTEMS

Dimitar Z. Dimitrov

**Abstract:** A system, part of a website building system, implemented on a server, the server having at least one processor and operating memory, including a system for image cropping, a system for data management working in parallel with the currently described system. The described system for optimization of images in a website building system solves several general problems. The system is a part of a complex solution of two major problems in the online visual editing systems, in particular the usage of too many resources, which in term leads to slow operations in the editing process and in terms bad user experience. The second problem is the high traffic generated by the website building systems. There are no publications of a similar research in the field of online visual editing tools or website builders in particular.

**Keywords:** browser rendering optimization, traffic optimization, resource optimization, image optimization, website building systems, website builders

## 1. Introduction in the field of research

The paper is focusing in the field of website building systems, multiwebsite systems and online visual editing tools in particular. The optimization presented by this paper is also related to the field of UX (User Experience) and the optimization of the browser client side. In the field of online applications this is related to all processes that are happening locally in the web browser of the user.

The modern systems for interactive website building are used for the creation of dynamic websites. The term “Multiwebsite system” [1] is related to a system for interactive website building, that can manage more than one website at the same time by a single user. The system for optimization of images is applicable to the multiwebsite systems. The website building systems can be a stand-alone application or be embedded in a bigger system. They can also be implemented for online, offline or partially online environment. The currently proposed system is applicable for online website building systems, e.g. Software as a Service (SaaS) based, implemented as an online application, working in a browser.

A website building system usually enables the user to create and edit websites using an application or a group of applications. The visually architected application consists of pages, containers and components. The pages usually are represented separately and contains components. The components usually are ordered in hierarchy in containers, which in term are ordered in the pages autonomously [2]. The components usually have attributes (size, position, color, border, type, etc.) and content (e.g. the path to the image in an image component). The pages can use templates, which can be templates of a whole page and templates of a component.

The components can be fixed or can be managed by a CMS, e.g. a component containing an image, selected from a database of available images. Such a database in this example can be extended by the user by uploading additional images. The components can also be based on entirely user driven content, e.g. a text component in which the user edits the text, or can be based on external content, static or dynamic, e.g. RSS reader, showing the information of an external data source or the content of a specific internet page.

The visual representation of a component can be modified by resizing, moving, rotating and other similar operations. It can also be modified by a set of visual settings and animations. A certain set of attributes (e.g. the button has a shape, color, text font) is defined as a component configuration. The configuration and the content of a component must be distinguishable [3].



## 2. Summary

The modern online website building systems are usually SPA (Single Page Applications), meaning application working dynamically on a single loaded web page. When a website building system is created as a SPA, numerous problems that needs solving arise. The presented paper proposes a solution for improving the UX of the visual editing client part of the website building systems.

When having a numerous visual elements, animations and effects, functioning in the system, even the modern browsers experience difficulties to operate smoothly in visual rendering and executing of programming code, which in terms executes the animation and the different effects. In addition to all this are added animation-heavy styles, requiring optimal usage of resources.

In the moment of writing this article, several browsers were identified, that were performing in visual rendering much better than the others. Nevertheless, in unoptimized website building systems, even those browsers create bad UX. It is expressed in skipping effects, slow work, browser blocking while processing heavy operations, an overall user feeling an experience of a heavy system. From the most widely used browsers, Internet Explorer 11 and recent versions of Firefox performed weakly in this regard. While the better working browsers were Chrome, Safari (without the Windows version), Opera and the latest version of Firefox with certain animation operation still rendering slowly. Despite the better performance of the latter, when the system misses a certain set of optimizations, the browsers show that they have not reached the level of development, that can provide a UX for a native application.

In terms of visual rendering, several main directions for optimization were identified for the application to operate while the system is scaling up. The currently described system proposes a solution for one of the identified directions for optimization.

In the visual editing system of a website building system the user works with multiple images and usually all images of a web page are rendered at the same time with the options for many operations of each of them. Operations like changing the image, styling, positioning, resizing, rotating, etc.

The currently described system proposes an approach for optimizing images, which in terms optimizes the so-called drawing of regions in the browser while executing the different operations, animations and effects.

### **Requirements for the image cropping system in the website building system**

An additionally implemented standard image cropping system is required for the needs of the currently described system. There are numerous libraries, online API services, also described methods and architectures described for implementing such an image cropping system. The main requirement for the image cropping system is to allow dynamic HTTP GET requests with dimension parameters. E.g. <https://websitebuildersystem.com/cropimage/w/100/h/100/mode/center>, which responds with the cropped image as a result, caching the image in parallel of the request for optimization purposes.

## 3. Related research in the field

In respect of the niche application of the proposed research, the related research is in the field of related add-on technologies, rather than the exact field itself, like numerous methods for image cropping, to help the add-on cropping system, required to work alongside the currently proposed system. The image cropping system is a major and very resource-heavy part of a website building system. The proposed system is an optimized approach for indirect (managed) usage of the cropping system.

An excellent approach for an image cropping method is the content-aware cropping [4]. This approach uses a machine learning framework, to identify the relative meaning of the regions of an image and to identify the important regions in regards of content, which must be included in the image after the cropping.

Another approach in image cropping is by using retargeting for controlled cropping [5]. This approach similarly to the previous one identifies the important regions of an image, and in addition calculates the possibility to include most of them in the result. It also calculates the weight and the importance for the image of each region separately. After an image is requested in several sizes, it will add the regions one by one, ordered by weight, and will crop the regions that are relatively not important to fit the requested size [6, 7].

The development of this approach can even result in changing entirely the original image, by recalculating and cropping the internal regions of the image:



**Fig 5.** Represents function realism instead of photo realism by image retargeting [8]

Further describing Figure 1, the input image consists of 3 regions of high importance, the two boys and the ball. The image is retargeted to be shown in smaller sizes. The retargeting algorithm manages to keep the relative positions of the important regions by including even their shadows. It is also possible for the cropping system a completely new version of the image, which does not keep the relative positions of the objects. This is an interesting approach in which each important object is important separately for itself, in this manner allowing the generation of diversified set of layouts and variants [9, 10].

#### **4. Full description**

In the development process of a standard website, the images are exported using a desktop graphic editing software for web usage in an exact size, which is the optimal size for the exact website. In the online systems for visual editing this image needs to be in high quality. E.g. if we want to resize the image to 3000px width, it means that the image must be loaded in at least that size if we want to bypass the other bad UX – a pixelized image. This in terms means that perhaps hundreds of high-quality images must be loaded in the web browser simultaneously. All this visual information needs to be rendered and redrawn by the browser for each operation that the browser executes. E.g. when an image element is being dragged on top of a gallery of 100 images. In this moment the browser must redraw the regions in which the image element is being dragged, which becomes a very intensive operation when all images are in high quality.

A seemingly possible solution can be reducing the image sizes by using one of the many known methods for image cropping. For best UX the image shouldn't be smaller than the container it is displayed in. In the other case it will result in quality reduction and again bad UX. Another seemingly possible solution is cropping to the exact container size, so that the image does not lose quality and in addition to be in a much smaller than the original size. In this case the browser will operate much more efficiently in redrawing of screen regions. In that moment comes the requirement for the image component to be able to be resized visually by the user. In this case e.g. from a size of 100px by 100px the user can resize the image to 3000px by 1000px, and the process

of resizing can be slow and smooth. In the case of exact container-size cropping, the cropping must be executed on every pixel in the process of resizing, while the container grows bigger than the image. In this case, to keep the image with good quality, not having to reduce the UX by showing the image pixelized, thousands expensive of cropping operations needs to be executed in a matter of seconds by a single user action. And all this needs to be transferred through the network to the user.

With these requirements, the initially seemingly possible solution is inapplicable. The required cost of resources is too high, comparing to the small goal that will be met. In addition to that, the browser requires a certain amount of resources for the rendering of the cropped images, and if this happens at a pixel precision it will again lead to bad UX and even to blocking the whole browser.

Many implemented website building systems use an approach that allows a certain level of bad quality of the images while resizing and crop the image when it reaches the final size. In the time of resizing, when the user wants to adjust the size of the image, he sees a low-quality image, not corresponding to the original quality.

This paper describes a system for image optimization in a website building system, which proposes a balanced approach for optimal usage of server resources, aiming to achieve optimal UX. In result, the user will use a light and fast website building system, images displayed in the visually highest possible quality in any size, even during the duration of the resizing operations.

$$\begin{bmatrix} 1500px & x & 1500px \\ 1000px & x & 1000px \\ 500px & x & 500px \\ 200px & x & 200px \end{bmatrix} \quad (1)$$

In (1) are represented regions of growth of the image during the resize operations. The parameters of the different growth regions can be optimized for each system individually, based on the most common use of the exact class of a website building system.

Depending on the processing methods in the implementation of the described system, the growth regions can be described in a two-dimensional array or as a matrix. The number of growth regions can be different for the different systems, based on the identified common use of images in the certain class of website building system. In this paper 4 growth regions are suggested as an optimal solution, based on a conducted research of most common uses in systems for building business information websites and online stores.

The growth regions also define the maximum number of possible crops of an image, having the requirement that the image cropping system will be caching each cropped size.

When the user uploads an image, e.g. with a file size of 8Mb, it will be cropped right away to the maximal size defined by the website building system. At that size the system will consider as the original image, which will be used for further cropping, while the uploaded file is not stored. To support the bigger screen resolution sizes, when a certain image needs to be a background on the whole width of the screen, the maximum resolution size at the time of writing this article should be about 3000px by 3000px. Experiments show that the size of this image if cropped in optimal quality will be around 1Mb. When this image is used for the entire background of a page, it is justified cost, but when we use a gallery of e.g. 100 images on a single page, while visually showing size of 200px by 200px, a more optimal solution must be implemented, allowing this gallery to resize smoothly while the user experiences high image quality, and in addition to bypass loading 100Mb of images in the visual editor. When we take the highest defined resolution as the original image, the growth regions are intermediate sizes, which the image can be further cropped into. When the component from Figure 2 is rendered in dimensions of 180x168 pixels, it is positioned in the smallest growth region, of the proposed regions in (1). In this example the image is loaded in 200x200px size. The following figures will demonstrate a steady resizing of the component, which will pass through the different growth regions. E.g. in the moment which the component reaches sizes above 201x201px, the image optimization system will trigger a background process of loading a bigger image

corresponding to the next growth region, which will be 500x500px, and in the moment when the background process completes, the loaded bigger image will replace the smaller image seamlessly for the user. Following this process, the image will keep its high quality in any region and in the same time be optimized so it doesn't cause the previously described issues.



**Fig. 6.** Represents an example image component in a website building system, which in its initial state has a width in the smallest growth region of up to 200px



**Fig. 7.** Demonstrates moving in a growth region of between 1000px and 1500px

When the component grows to the next growth region in Figure 3, the same operation for loading the image with the size of the following region repeats. In this case this is 1500x1500px. And when the component reaches a size outside of the last growth region, the system returns the original image, limited by the website building system definition of maximal size. This limitation depends entirely on the devices that the website builder system will support. If a system needs to support devices with 8K resolution, to provide best quality and UX, the defined max image width



must be above 8000px. At the time of writing the current paper 8K is still a resolution, which is rarely used for loading websites, but this is very possible to change in the future. At the time of writing the current paper, the most used resolutions are below 3000px width, which is defined as a very good definition of a maximal image size, which the system will be cropping all images to on upload.



**Fig. 8.** Demonstrates a wide screen setting of an image, which sets 100% screen width of the image element. This is a widely used possibility in responsive design

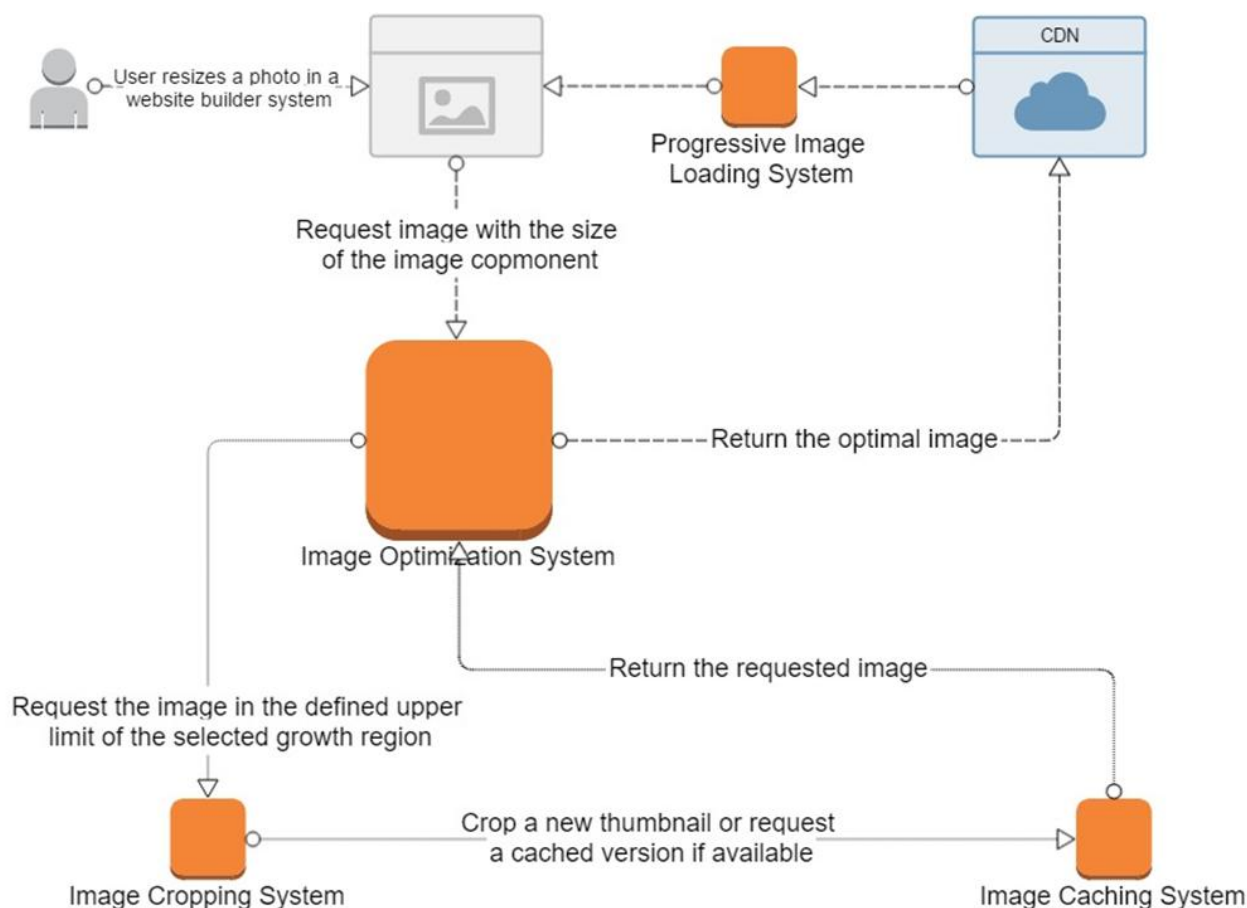
Figure 4 demonstrates the case which will visualize the original image, based on the maximum defined system size. In the experimental system the size is 3000x3000. When the image has a wide screen setting to 100% screen size, it will be with high quality, set to this maximum size. And quality reduction will be observed on devices with higher resolutions like 8K (Ultra HD). When a support for higher resolutions is required, the proposed system can be implemented with a higher number of growth regions, following the same principals. The regions are not required to be equally distant and big in size, but they need to follow the identified most widely used cases of image sizes of the class of website building system they are used for.

In this example, there can be a maximum of 5 possible crops of each image in the working process of a user, in this account including the defined maximum system size, cropped right after the image is uploaded. The cropping operation is executed only once per region, after which the image cropping system caches the cropped image. For all following requests it just responds with the cached images.

When considering the optimization of server resources and traffic, using standard approaches for visual editing, the images are loading in their original size, limited by the platform and in the time of the visual editing are shown in their original size, independently of the component size. In the proposed system an optimized image of 3000x3000px is cropped to around 1Mb (Varies depending on image). If we have an example of a single page having a gallery of 100 images, which are visualized 5 on a row with a size of up to 200px width of the image component, in the proposed system the images will be displayed in 200x200px size, leading to an image of about 15Kb (Tested with same images as above), and when resizing to the next region it reaches 80Kb, displaying in top quality on devices with up to 3000px resolution. In the given example, if we exclude all the other benefits for the UX, improving the browser performance and reducing server resources costs with minimal set of requests, the optimization of traffic and the load speed can be increased over 50 times. It must be considered that the example shows the optimization in 1 page with 1 gallery of

100 images, and website building systems can be used by millions of users simultaneously editing their websites. Currently there are website building systems with more than 100 million of users.

As further optimization to the proposed system, although not required, additional systems can be included. Like a system for progressive image loading [11] and a content delivery network (CDN) [12]. The system for progressive image loading is an intermediate system between the cropping system and the currently described system. When an image is requested by the currently described system, a very small image is generated around the size of 2Kb, which creates a visual idea of the original image. It sends initially the small image, which the user sees almost immediately, while the requested image is being transferred through the network. This is yet another layer that helps UX. A CDN network in addition can be created or can be rented as a service. One of the biggest CDN providers are CloudFlare and AWS CloudFront. Using a CDN all images (and static resources) are cached to geographically closest servers to the end user. This is yet another step, in improving user experience by allowing faster load speeds and in the same time reduce server load. In addition, we can mention that most ISPs provide a better local connectivity in comparison with the international one, which is another advantage to implement a CDN.



**Fig. 9.** An example workflow and architecture of a system for image optimization, an image cropping system, a system for progressive image loading and a CDN in a website building system

In addition to each underlying system, including the currently described system for image optimization, load balancing can be included to scale it. Each system separately can be executed on a dynamically growing set of servers using technologies for autoscaling. This way in any given moment, each system can be served by a growing set of servers, needed to meet the load at the exact moment. Many cloud service providers link the costs to the hourly use of resources, including the number of server instances. Using autoscaling bypasses the need of renting unneeded resources

and allows each server to be scaled to hundreds of instances in a short period of time, which after the peak can be scaled down again to one instance.

Each additional system in Figure 5 can be viewed as a closed system with an input and output, as all additional systems can be implemented using the numerous known methods and technologies. Many of them can be used as a service by different SaaS providers, usually offering public APIs for integration.

## 5. Conclusions

The modern online website building systems are resource-heavy in both the client side as an app working in the browser and the server and network side. The proposed system for image optimization in a website building system is a solution, which improves the browser performance and UX on the client side, to minimize traffic progressively and minimize the server resources needed. This leads to increasing the overall UX and minimizing the server and network costs. The system proposes a balanced approach in using the image cropping system, available in a website building system.

## References

- [1]. Dimitar Dimitrov, Standardization based on multiwebsite system integration – International Science and Technology Conference, Dubai, 2012
- [2]. R. Ben-Aharon, N. Abrahami, System and method for the creation and use of visually-diverse high-quality dynamic visual data structures, US 2016/0357527 A1 - United States Patent Application Publication, 2016
- [3]. A. Geva, O. Gartzman, G. Kaplan, System and method for the generation of an adaptive user interface in a website building system, US 2016/0357527 A1 - United States Patent Application Publication, 2016
- [4]. L. Kennedy, R. Zwol, N. Torzec, Learning Crop Regions for Content-Aware Generation of Thumbnail Images – ACM International Conference on Multimedia Retrieval, 2011
- [5]. R. Palp, P. Mitra, J. Mukherjee, Image Retargeting Using Controlled Shrinkage – Eighth Indian Conference on Computer Vision, Graphics and Image Processing, 2012
- [6]. B. Suh, H. Ling - Automatic Thumbnail Cropping and its Effectiveness – 16th annual ACM symposium on User interface software and technology, 2003
- [7]. Masashi N., T. Okabe, Sensation-based Photo Cropping – 17th ACM international conference on Multimedia, 2009
- [8]. V. Setlur, S. Takagi, Automatic Image Retargeting – SIGGRAPH, 2005
- [9]. J. Hu, S. Chen, Fast Image Rearrangement via Multi-Scale Patch Copying – 18th ACM international conference on Multimedia, 2010
- [10]. Amrutha S., Shylaja S., A Smart Automatic Thumbnail Cropping Based on Attention Driven Regions of Interest Extraction – 2 nd International Conference on Interaction Sciences: Information Technology, Culture and Human, 2009
- [11]. C. Harrison, A. Dey, S. Hudson, Evaluation of Progressive Image Loading Schemes – SIGCHI Conference on Human Factors in Computing Systems, 2010
- [12]. P. Ganti, Real World Content Delivery Networks: Architecting Global Applications - O'Reilly Media, 2017

### For contacts:

PhD candidate Dimitar Z. Dimitrov  
Technical University of Varna, Bulgaria  
E-mail: dimitrov@oxxy.com

# КЛАСИФИЦИРАНЕ НА МНЕНИЯ ОТ ПОТРЕБИТЕЛСКИ КОМЕНТАРИ НА АНГЛИЙСКИ ЕЗИК ЧРЕЗ АЛГОРИТМИ НА МАШИННОТО ОБУЧЕНИЕ

Нели Ан. Арабаджиева – Калчева, Недялко Н. Николов

**Резюме:** Целта на публикацията е сравнение на точността на алгоритми на машинното обучение при класифициране на потребителски мнения на английски език. Представен е нов алгоритъм, базиран на Наивния Бейсов класификатор, използващ Лапласово разпределение и именуван Лапласов Наивен Бейсов класификатор. Анализираните алгоритми са: Лапласов Наивен Бейсов класификатор, Гаусов Наивен Бейсов класификатор, Бернулиев Наивен Бейсов класификатор, Мултиноминален Наивен Бейсов класификатор, Метод на най-близък съсед, Метод на опорните вектори, Дърво на решението, ансамблов метод Случайна гора и ансамблов метод Адаптивен усилвател.

**Ключови думи:** Наивен Бейсов класификатор, Метод на най-близък съсед, Метод на опорните вектори, Дърво на решението, Случайна гора, AdaBoost, класификация на текст, машинно обучение

## Classification of user comments in english through machine learning

Neli An. Arabadziewa – Kalcheva, Nedyalko N. Nikolov

**Abstract:** The purpose of the publication is to compare the accuracy of machine learning algorithms to classify user comments in English. A new algorithm is introduced, based on the Naive Bayes Classifier using Laplace Distribution and a named Laplace Naive Bayes Classifier. Analyzed algorithms are: Laplace's Naive Bayes classifier, Gaussian Naive Bayes classifier, Bernoulli Naive Bayes classifier, Multinomial Naive Bayes classifier, K-nearest neighbours, Support Vector Machine, Decision trees, Random Forest and AdaBoost.

**Keywords:** Naive Bayes classifier, K-nearest neighbours, Support Vector Machines, Decision trees, Random Forrest, AdaBoost, text classification, machine learning

## 1. Увод

Масовото използване на социални медии предостави средства на потребителите да публикуват мнения, отзиви и оценки на стоки и услуги. Нарасналото съдържание в различните коментари доведе до невъзможност на тяхното ръчно обработване. Използването на алгоритми на машинното обучение за класификация на текст позволява да се съкрати времето за проучване и анализ на потребителското мнение.

Задачата за класификация на текст се състои в определяне на принадлежност на текста към даден клас въз основа на анализ на набор от признаци, характеризиращи дадения текст.

## 2. Изложение

### 2.1. Наивен Бейсов класификатор

Един от класическите алгоритми в машинното обучение е Наивният Бейсов класификатор, който се базира на теоремата на Бейс за определяне на апостериорната вероятност за настъпване на дадено събитие.

Теорема на Бейс

$$P(y = c|x) = \frac{P(x|y=c)P(y=c)}{P(x)} \quad (1)$$



където:  $P(y=c|x)$  е вероятността обект да принадлежи на клас  $c$  (апостериорна вероятност на класа)

$P(x|y=c)$  – вероятността обектът  $x$  да се среща в средата на обекта на класа  $c$

$P(y=c)$  – безусловна вероятност да се среща обект  $y$  в клас  $c$  (априорна вероятност на класа)

$P(x)$  – безусловна вероятност на обекта  $x$

Целта на класификацията се състои в това, да се определи към какъв клас принадлежи обектът  $x$ . Следователно е необходимо да се намери вероятностен клас на обекта  $x$ , т.е. необходимо е от всички класове да се избере този, който дава максимална вероятност  $P(y=c|x)$ .

$$c_{opt} = \arg \max_{c \in C} P(x|y=c)P(y=c) \quad (2)$$

Известни алгоритми от типа Наивен Бейсов класификатор са: Бернулиев, Мултиноминален и Гаусов, свързани с различни предположения за разпределението на признаците.

## 2.2. Бернулиев Наивен Бейсов класификатор

Бернулиевият Наивен Бейсов класификатор представя документа като бинарен вектор на пространството. Стойност единица показва дали думата се появява поне веднъж в документа. С такова представяне се прави наивното предположение на Бейс, че вероятността всяка дума да се среща в текста е независима от появата на други думи.

Нека  $x \in [0,1]$  е бинарна променлива, която има вероятност за поява  $\theta$ .

Тогава:

$$P(x|\theta) = \begin{cases} \theta & , \text{ ако } x = 1 \\ 1 - \theta & , \text{ ако } x = 0 \end{cases} \quad (3)$$

## 2.3. Мултиноминален Наивен Бейсов класификатор

Мултиноминалният Наивен Бейсов класификатор приема предположението, че признаците са разпределени мултиноминално.

Нека  $x_i \in \{1, \dots, K\}$ , имат вероятности за поява  $\theta_1, \dots, \theta_K$  [7].

Тогава вероятността за поява на събитието  $x$  при даден признак  $\theta$  е:

$$P(x|\theta) = \frac{n!}{x_1! \dots x_K!} \prod_{i=1}^K \theta_i^{x_i} \quad (4)$$

$$\text{където: } n = \sum_{i=1}^K x_i \quad (5)$$

Мултиноминалният Наивен Бейсов класификатор изчислява честотата на поява на всяка дума в документите.

## 2.4. Гаусов Наивен Бейсов класификатор

Гусовият Наивен Бейсов класификатор се основава на предположението, че признаците са с нормално разпределение.

Функцията на вероятностната плътност на нормално разпределена случайна величина е:

$$P(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (6)$$

където:

$\mu$  - средната стойност;

$$\mu = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2 \quad (7)$$

$\sigma$  - стандартното отклонение

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2 \quad (8)$$

## 2.5. Наивен Бейсов класификатор, използващ Лапласово разпределение

Наивен Бейсов класификатор, използващ Лапласово разпределение, е нов алгоритъм, представен от авторите, базиран на Наивния Бейсов класификатор, използващ Лапласово разпределение и именуван Лапласов Наивен Бейсов класификатор.

Класическото разпределение на Лаплас се определя като функция за вероятностна плътност – probability density function (pdf). [4]

$$f(x, \theta, \phi) = \frac{1}{2\phi} \exp\left(-\frac{|x - \theta|}{\phi}\right), \quad -\infty < x < \infty \quad (9)$$

където:  $\theta \in (-\infty, \infty)$  и  $\phi > 0$  са локален и скален параметър респективно.

Това е разпределение на вероятностите, което се максимизира, когато локалният параметър е медиана. Лапласовото разпределение е симетрично разпределение, чиито опашки са по-малко остри от Гаусовото разпределение.

## 2.6. Метод на най-близкия съсед – KNN (K-nearest neighbours)

Методът на K най-близкия съсед е алгоритъм за класификация на обекти, изчисляващ дистанцията между всеки два обекта от обучаващата извадка, чрез използване на подходяща функция за измерване на разстояние между две точки. Класът на класифицирания обект – това е класът на новия обект, намерен въз основа на заключенията, направени по отношение на K най-близки съседи.

Функция, изчисляваща Евклидово разстояние:

$$\rho(x_i, x_j) = \sqrt{\sum_{k=1}^m w_k (x_i^{(k)} - x_j^{(k)})^2} \quad (10)$$

където:

$x_i = (x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(m)})$  – вектор от m-признаци на i- я обект

$x_j = (x_j^{(1)}, x_j^{(2)}, \dots, x_j^{(m)})$  – вектор от m-признаци на j- я обект

Други известни функции, изчисляващи разстояние между две точки, са:  $L_p$  – метрика,  $L_\infty$  – метрика,  $L_1$  – метрика, Ланс – Уилямс Косинусова мяра.

## 2.7. Метод на опорните вектори - SVM (Support Vector Machine)

Методът на опорните вектори (SVM, Support Vector Machines) представя обучаващите примери като точки в n-мерно пространство. Примерите са проектират в пространството по такъв начин, че да бъдат линейно разделими. При работа с два класа се търси начин да се начертае линия, която да разделя данните от двата класа. Линията, която разделя данните, се нарича разделителна хипер равнина. Тази хипер равнина трябва да се избере по такъв начин, че да се намира възможно най-далеч от примерите и на двата класа.

Функцията  $f(x)$  на линейната класификация е във вида: [5]

$$f(x) = w^T x + b \quad (11)$$

където:  $w^T$  е тегловен вектор, а  $b$  е отклонението.

Целта е да се намерят стойностите на  $w^T$  и  $b$ , които ще определят класификатора. За да се направи това, е необходимо да се намерят точките с най-малко отклонение, което трябва да се максимизира.

## 2.8. Алгоритъм Дървета на решенията (Decision trees)

Алгоритъмът Дървета на решенията строи класификационни дървета за изграждане на модел въз основа на статистически оценки и поредици от проверки на логически условия.

За всеки възел на дървото се избира оптимално правило за разделяне, включващо избор на атрибут и стойност на този атрибут, по който да стане разделянето. Изборът на атрибут се основава на информационната печалба при разделяне на обектите по съответния атрибут.

Информационната печалба на атрибута  $A$  по отношение на набор от примери  $S$  се изчислява като: [6]

$$Gain(S, A) = E(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} E(S_v) \quad (12)$$

където:  $S_v$  е множество от примери, които имат стойност  $v$  за атрибута  $A$ ,  $E(S)$  - ентропия на  $S$  множество от примери.

Ентропията се представя в следния вид:

$$E(S) = \sum_{i=1}^n -p_i \log_2(p_i) \quad (13)$$

където:  $n$  – брой примери,  $p_i$  – вероятността да се избере  $i$ -клас.

Алгоритъмът CART (Classification and Regression) използва индекс Gini, който оценява разстоянието между разпределените класове:

$$Gini(S) = 1 - \sum_{i=1}^n p_i^2 \quad (14)$$

## 2.9. Ансамблов алгоритъм Случайна гора (Random Forest)

Случайна гора класификаторът е ансамблов алгоритъм. Ансамбловите алгоритми съчетават повече от един алгоритъм от един и същи или различен алгоритъм за класификация.

Даден класификатор  $f_m$  ( $m=1,2,\dots,M$ ) се обучава на базата на повторна извадка, като резултантната функция е функция на “гласуването”[1].

$$f(x) = \sum_{m=1}^M \frac{1}{M} f_m(x) \quad (15)$$

където  $f_m$  е  $m$ -тото дърво. Тази техника се нарича bagging (bootstrap aggregation). В общия случай примерите се подбират така, че всеки екземпляр има еднакъв шанс да попадне в обучаващата извадка.

RandomForest класификаторът се различава от общата схема, като създава набор от дървета за вземане на решения от произволно избран набор от тренировъчни комплекти. Алгоритъмът използва усредняване с цел подобряване точността на прогнозирането. Всяко отделно дърво представлява случайна извадка, като изборът е независим и използва еднакво разпределение на всички дървета. Всяко дърво осъществява класификация, след което статистическата мода на всички отделни класификации се взема като крайна класификация на гората.

## 2.10. Ансамблов алгоритъм Адаптивен усилвател (AdaBoost)

Алгоритъмът Адаптивен усилвател е представител на ансамбловия алгоритъм boosting. Това е итерационен алгоритъм, при който „силният“ класификатор допуска малки грешки при обучението на основата на „слабия“ класификатор, който правилно класифицира повече от 50%. В ансамбловия алгоритъм Adaboost различните класификатори се обучават последователно. Всеки нов класификатор е обучен въз основа на ефективността на вече съществуващи обучени. Усилването се състои в съсредоточаване на новите класификатори върху данни, които са били погрешно класифицирани от предишни класификатори.

За да получи отговор от всички слаби класификатори, алгоритъмът Adaboost определя стойности  $\alpha$  към всеки един от класификаторите. Стойностите  $\alpha$  се основават на грешката на всеки слаб класификатор.

Грешката  $\varepsilon$  се изчислява със следната формула:

$$\varepsilon = \frac{\text{брой некоректно класифицирани екземпляри}}{\text{общ брой екземпляри}} \quad (16)$$

$\alpha$  се изчислява чрез  $\varepsilon$

$$\alpha = \frac{1}{2} \ln \left( \frac{1 - \varepsilon}{\varepsilon} \right) \quad (17)$$

### 3. Оценка на точността на класификацията

В процеса на класификация на текст данните се разделят на две части – едната част се нарича обучаваща извадка (training-set), която се използва за обучение на модела, и тестова извадка (test-set), използвана за тестване на модела. Обучителните данни „обучават“ тестовите данни.

За оценяване на класификатора се използва мярката точност (Accuracy), която се изчислява като отношението на правилно класифицирани примери към общия брой обекти от тестово множество. За получаване на по-устойчива оценка се прилага методът на кръстосаното валидиране (cross-validation), при който се гарантира равен брой участия на всеки обект в обучаващата извадка и точно едно участие в тестовата извадка.

### 4. Изследване, получени резултати и анализ

В изследването алгоритмите на машинното обучение са реализирани на езика Python. Използвани са две външни библиотеки – nltk (<http://www.nltk.org>) и scikit-learn (<http://scikit-learn.org>). За оценка на класификацията е използвано 10-кратно кръстосано валидиране и мярка точност.

Анализираните алгоритми са: Лапласов Наивен Бейсов класификатор, Гаусов Наивен Бейсов класификатор, Бернулиев Наивен Бейсов класификатор, Мултиноминален Наивен Бейсов класификатор, Метод на най-близък съсед, Метод на опорните вектори, Дърво на решението, ансамблов метод Случайна гора и ансамблов метод Адаптивен усилвател. Основна причина за избора на тези алгоритми е тяхната различна природа, т.е. различният начин на разделяне на съвкупността от данни на отделни класове.

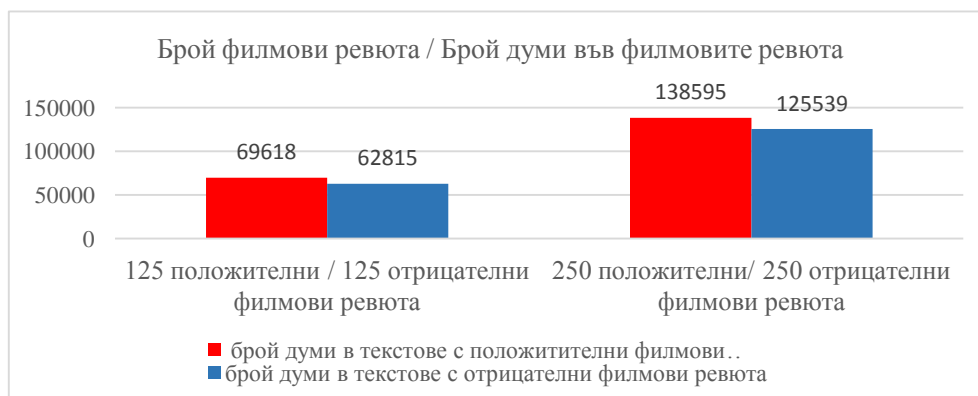
Целта на изследването е да се провери кой от алгоритмите класифицира с най-висока точност при различен обем от данни.

В доклада са въведени следните съкращения на използваните алгоритми в текста:

- Лапласов Наивен Бейсов класификатор – LNB
- Гаусов Наивен Бейсов класификатор – GNB
- Бернулиев Наивен Бейсов класификатор – BNB
- Мултиноминален Наивен Бейсов класификатор – MNB
- Метод на най-близък съсед – KNN (K-nearestneighbours)
- Метод на опорните вектори – SVM (Support Vector Mashine).
- Дърво на решението – DT ((Decision trees)
- Ансамблов метод Случайна гора – RF ((Random Forest)
- Ансамблов метод Адаптивен усилвател - AB (AdaBoost)

За изследването е избран набор от данни, състоящ се от 500 мнения на потребителски коментари, представляващи филмови ревята на английски език, от които 250 положителни, съдържащи 138 595 думи, и 250 отрицателни филмови ревята, съставени от 125 539 думи. Множеството от данни е намалено двойно и така е получен нов набор от 250 филмови ревята, от които 125 положителни, съдържащи 69 618 думи, и 125 отрицателни филмови ревята, съставени от 62 815 думи. Данните са взети от стандартния модул NLTK – Natural

Language Toolkit, който е open source библиотека на Python. Филмовите ревюта са предоставени през 2004 година от Bo Pang и Lillian Lee (Department of Computer Science, Cornell University, Ithaca, NY).

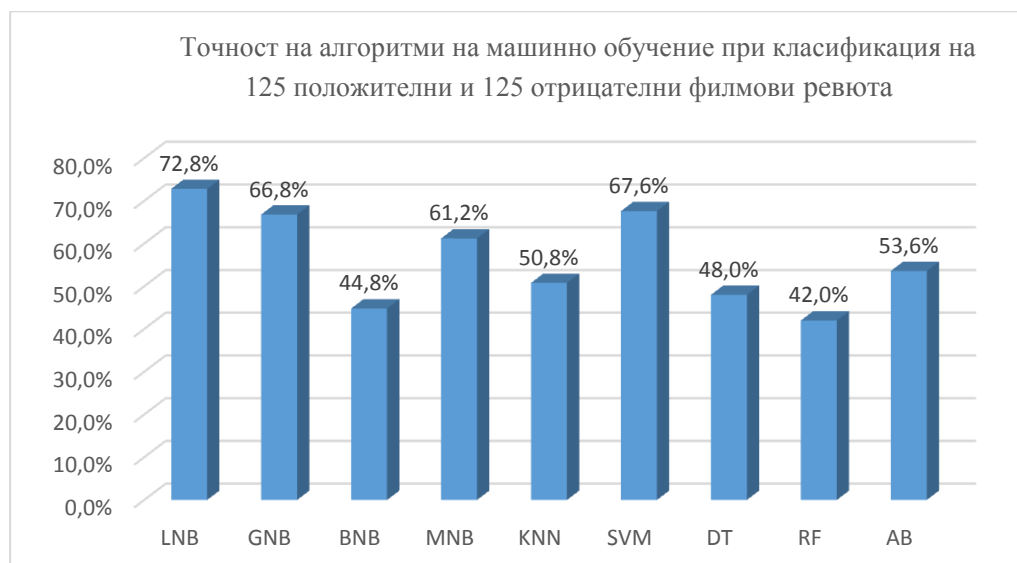


**Фиг. 1.** Разпределение на брой думи в текстове с положителни и отрицателни филмови ревюта

Видно от фигура 1, броят думи в положителните филмови ревюта и в двата набора от данни е повече от броя думи в отрицателните филмови ревюта, като с нарастване на броя отзиви разликата също се увеличава.

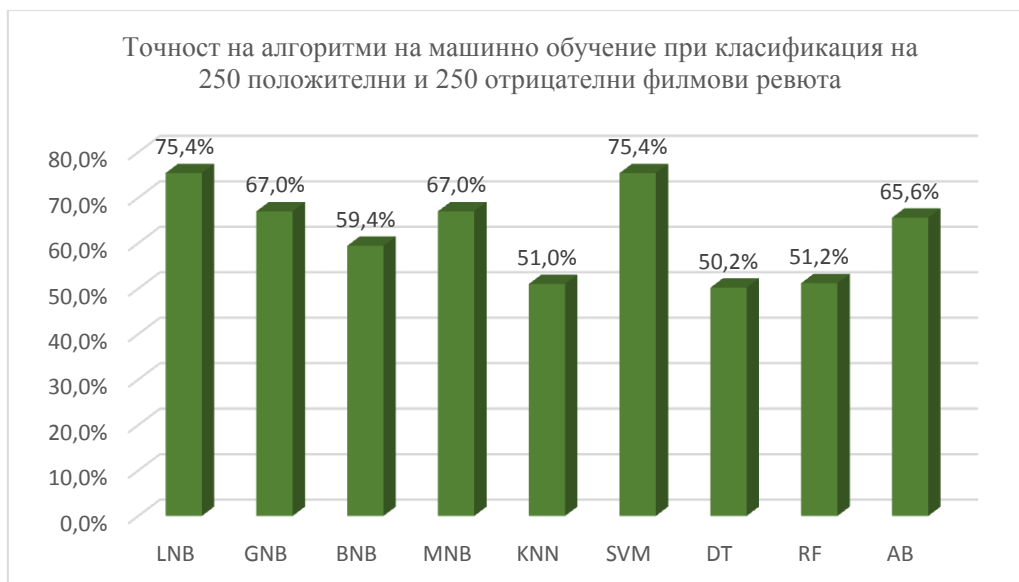
Първоначално са класифицирани 125 отрицателни и 125 положителни филмови ревюта на английски език.

Резултатите показват (фигура 2), че с най-висока точност е класифицирал Лапласовият Наивен Бейсов класификатор с преднина от следващия го алгоритъм от  $\approx 5.2\%$ . Вторият по точност е Методът на опорните вектори (с линейно ядро), следван от Гаусовия Наивен Бейсов класификатор, с малка разлика помежду им от  $0.8\%$ . Мултиномиалният Наивен Бейсов класификатор е на четвърта позиция и е със стойност по-висока от средната, която е  $\approx 11.6\%$  по-малка от най-добре представилия се и с  $\approx 19.2\%$  по-голяма от последния алгоритъм в класацията. Ансамбловият алгоритъм Адаптивен усилвател класифицира филмовите ревюта със средна точност спрямо всички класификатори. Бернулевият Наивен Бейсов класификатор, Методът на най-близкия съсед и алгоритъмът Дърво на решенията са с резултати по-ниски от средните. С най-ниска точност е класифицирал алгоритъмът Случайна гора, който, в сравнение с първенеца, има разлика от  $\approx 30.8\%$ .



**Фиг. 2.** Класификация на 125 положителни и 125 отрицателни филмови ревюта чрез алгоритми на машинното обучение

При двойно увеличаване както на положителните, така и на отрицателните коментари от 125 на 250, т.е. сумарно 500 филмови ревюта, всички алгоритми увеличават точността си (фигура 3).



**Фиг. 3.** Класификация на 250 положителни и 250 отрицателни филмови ревюта чрез алгоритми на машинното обучение

Лапласовият Наивен Бейсов класификатор и Методът на опорните вектори (с линейно ядро) са с равни и най-високи резултати. На втора позиция са Гаусовият Наивен Бейсов класификатор и Мултиноминалният Наивен Бейсов класификатор с разлика спрямо първите от 8.4%. Методът на опорните вектори (с линейно ядро) е разпознал коментарите със 7.8% по-точно спрямо предходното изследване, докато при Лапласовия Наивен Бейсов класификатор нарастването е  $\approx 2.6\%$ . Мултиноминалният Наивен Бейсов класификатор е увеличил точността си с  $\approx 5.8\%$ , докато при Гаусовия Наивен Бейсов класификатор увеличението е само 0.2%. Ансамбловият алгоритъм AdaBoost при нарастване на данните класифицира със значителна преднина от предишното изследване от 12%. С най-голяма разлика в положителна посока е Бернулевият Наивен Бейсов класификатор – 14.6%. Алгоритъмът Случайна гора във второто изследване вече не е на последна позиция и е увеличил точността си със значителните 9.2%. Алгоритъмът Метод на най-близкия съсед при двойно повече филмови ревюта е определил вярно коментарите с незначителна преднина от  $\approx 0.2\%$ . С най-ниска точност е класифицирал алгоритъмът Дърво на решенията, с разлика от предходното изследване от  $\approx 2.2\%$ .

## 5. Заключение

Проведените изследвания показват, че при класификация на потребителски мнения на английски език точността на представения алгоритъм, базиран на Наивния Бейсов класификатор, използващ Лапласово разпределение и именуван Лапласов Наивен Бейсов класификатор, е най-висока и е съизмерима с точността на Метода на опорните вектори. Всички алгоритми, използвани в сравнението: Лапласов Наивен Бейсов класификатор, Гаусов Наивен Бейсов класификатор, Бернулиев Наивен Бейсов класификатор, Мултиноминален Наивен Бейсов класификатор, Метод на най-близък съсед, Метод на опорните вектори, Дърво на решението, ансамблов метод Случайна гора и ансамблов метод Адаптивен усилвател увеличават точността си при увеличаване обема на данните.

## Литература

- [1]. Золотых, Н. Ю. Машинное обучение и анализ данных (Machine Learning and Data Mining)
- [2]. Уиткин Л. В. Машинное обучение (Machine Learning) Метрические методы классификации и регрессии, 2017
- [3]. Harrington, P., Machine Learning in Action, 2012
- [4]. Kozubowski, T.K. Podgorski, The Laplace Distribution and Generalizations, 2014
- [5]. McCallum, A., K. Nigam, A comparison of event models for Naive Bayes text classification. Papers from the 1998 AAAI Workshop, 1998, pp. 41-48.
- [6]. Mitchell, T. Machine Learning, McGraw-Hill Science/Engineering/Math, March 1, 1997
- [7]. Murphy K.P. Machine Learning A Probabilistic Perspective, 2012

### За контакти:

ас. инж. Нели Ананиева Арабаджиева - Калчева  
катедра „Софтуерни и интернет технологии”  
Технически университет - Варна  
E-mail: [n\\_kalcheva@tu-varna.bg](mailto:n_kalcheva@tu-varna.bg)

доц. д-р инж. Недялко Николаев Николов  
катедра „Софтуерни и интернет технологии”  
Технически университет - Варна  
E-mail: [ned.nikolov@tu-varna.bg](mailto:ned.nikolov@tu-varna.bg)



**СОФТУЕРНИ И ИНТЕРНЕТ  
ТЕХНОЛОГИИ**

# ИЗСЛЕДВАНЕ НА ПРОЛОЖИМОСТТА НА СЪЩЕСТВУВАЩИ МЕРКИ ЗА СХОДСТВО ЗА АНАЛИЗ НА ПОДОБИЕ МЕЖДУ ИЗОБРАЖЕНИЯ

Мирослав Б. Маринов, Ирена М. Въллова, Йордан И. Калмуков

**Резюме:** Този доклад представя подробен анализ на приложимостта на някои от най-популярните мерки за сходство за изчисляване на степента на подобие между изображения. Всички мерки са експериментално изследвани и е направен сравнителен анализ между тях. Използвани са 2 множества от изображения. Едното се състои от един и същ или много подобни обекти, снимани от различни ъгли, а второто – от произволно подбрани изображения, включващи множество различни обекти и различни цветови схеми и текстури.

**Ключови думи:** мерки за сходство, дивергенция, снимки, множества от изображения, обекти.

## Comparative analysis of existing similarity measures employed for content-based image retrieval

Miroslav B. Marinov, Irena M. Valova, Yordan I. Kalmukov

**Abstract:** This paper presents an in-depth comparative analysis of some of the most commonly used similarity measures in the context of content-based image retrieval. All metrics are experimentally evaluated and compared to real human judgements for perception of similarity. To increase objectivity of evaluation we used two different data sets – one composed of the same or very similar objects shot from different view angles and another one that consists of randomly selected images (including different objects and colour schemes).

**Keywords:** similarity measures, divergence, pictures, sets of images, objects.

### 1. Проблем на изследването

Търсенето измежду милиони или милиарди изображения е сложна и времеемка операция, но едновременно с това е и най-важната и желана функционална възможност на базите от данни от изображения. Съвременните търсещи машини и социални медии изключително много разчитат на съществуващите методи за анализ на съдържанието на изображения, за да могат правилно да обработят, категоризират и индексират милионите снимки, качвани всеки ден от техните потребители. За да бъде смислено и полезно обаче, търсенето трябва да бъде сравнително точно и да се извършва за разумен интервал от време. Тъй като повечето социални медии и системи за търсене са реализирани като уеб приложения, е желателно всички процедури по търсене да приключват в рамките на 2-3 секунди или потребителите просто може да не изчакат резултатите и да затворят браузъра преждевременно.

За да се отговори на изискванията за бързодействие, е необходимо от изображенията да се извлекат подходящи метаданни, които да бъдат индексирани и съхранени по начин такъв, че да позволи бързата им обработка в последствие. Самото търсене може да се реализира чрез сравнение на метаданните, извлечени от изображението заявка (предоставяно от потребител или външно приложение) с метаданните на всички изображения, съхранени в БД. Изключително важно е при представяне на резултатите от търсенето те да бъдат сортирани по намаляващ ред на степента им на подобие със заявката. Степента на подобие обикновено се изчислява чрез подходяща мярка за сходство, като тя се избира в зависимост от типа на метаданните. Именно избраната мярка за сходство определя както точността на търсене, така



и времето за изпълнение. Изчисляването на степента на подобие между изображенията представлява най-често повтаряната (а от там и най-времеемката) операция в цялата система, тъй като при всяка една потребителска заявка е необходимо да се изчислят подобията с всички индексирани в системата изображения. Затова и изборът на подходяща мярка за сходство е ключов (основен) при разработването на всяка система за анализ и търсене на изображения.

## 2. Описание на изследването

В тази статия са разгледани мерки за сходство, подходящи за определяне на степента на подобие между изображенията. В следващите подраздели последователно са описани преобразуванията за извличане на метаданни от изображенията; изборът на тестови извадки от изображения; избраните мерки за сходство, които се оценяват; показани са резултатите; и са направени изводи.

### 2.1. Изображения за изследване

Използваните изображения са с размерност 128 x 128 пиксела. С цел оптимизиране на времето за търсене и по-бързата им обработка, изображенията са разделени на 32x32 области, като за всяка от тях се съхранява само доминантният цвят в нея. Благодарение на това системата може да работи с изображения с различни резолюции – независимо от оригиналните размери на изображенията, вътрешно те се представят като матрици от 32 x 32 елемента. За откриване на подобни изображения цветовете в тях се редуцират до 64, като номерът/индексът на цвета (от 1 до 64) се използва и като негов код.00

Важно е да се отчете, че използваните изображения са взети от 0 и представляват 100 обекта, които са снимани под различен ъгъл. От тях са генерирани 2 множества от изображения, които представляват съответно множество от близки по цвят 3 обекта (домат, ягода и камба) и множество от различни по цвят и ориентация обекти (около 100), които са произволно избрани. Трябва да се вземе под внимание, че е напълно възможно една и съща снимка да участва и в двете множества от изображения. И двете множества съдържат по 216 изображения.

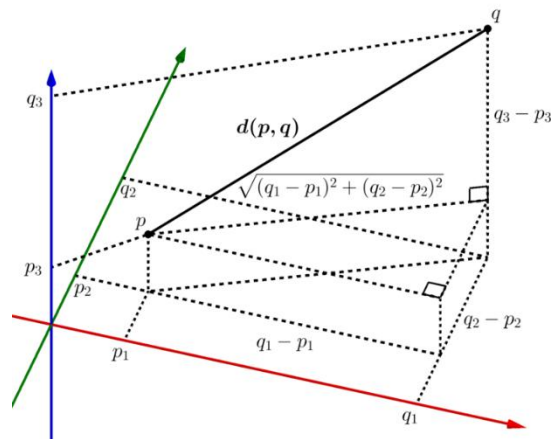
### 2.2. Мерки за сходство

За да се оцени степента на подобие между избрано изображение (заявка) с всички останали изображения (например в една база от данни от изображения) е много важно да се избере какъв да бъде критерият за оценяване на подобие. При сравняването на изображения не е възможно да се използват стандартни методи за търсене в бази от данни и това налага да се използват критерии за оценка на близостта между множества от елементи/обекти в многомерно пространство.

Съществуват различни такива мерки, като това изследване касае следните:

- **EUCLIDEAN DISTANCE**–Евклидовото разстояние (формула (1)) се изразява в намиране на най-краткото разстояние между две точки. То е винаги положителна стойност, поради което посоката няма значение, и е 0 само и единствено когато двете точки съвпадат. Пример може да се види на фигура 1.0

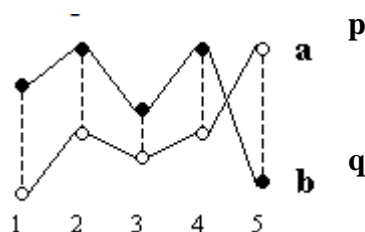
$$d(p, q) = d(q, p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (1)$$



Фиг. 1. Евклидово разстояние

- **CITYBLOCK DISTANCE** – Известно е още като разстояние в Манхатан (формула (2)) и се изразява в намирането на разстояние между две точки като се сумира хоризонталното и вертикалното разстояние между точките по координатната система. Пример може да се види на Фигура 2.0

$$d(q, p) = \sum_{i=1}^n |q_i - p_i| \quad (2)$$



Фиг. 2. Разстояние в Манхатън

- **CHI-SQUARE DISSIMILARITY** – Тази мярка се базира на чи-квадратния тест за равенство между две множества (формула (3)). Използва се по подразбиране за бройни данни.0

$$d(q, p) = \sum_{i=1}^n \frac{(q_i - \mu_i)^2}{\mu_i} \quad (3)$$

- **KULLBACK-LEIBLER DIVERGENCE** – дивергенцията на Kullback-Leibler (наричана още относителна ентропия) е мярка за това как едно разпределение на вероятностите се различава от второ, референтно разпределение на вероятностите (формула (4)). Това е асиметрична мярка за разпространение и по този начин не се квалифицира като статистически показател за разпространение.0

$$d(q, p) = \sum_{i=1}^n q_i \log \frac{q_i}{p_i} \quad (4)$$

- **JEFFREY DIVERGENCE** – дивергенцията на Jeffrey (формула (5)) е мярка за сходство, която показва как едно разпределение на вероятности се различава от второ, референтно разпределение на вероятности като се вземат предвид и двете позиции/индекси на сравняваните вектори.0

$$d(q, p) = \sum_{i=1}^n (q_i \log \frac{q_i}{\mu_i} + p_i \log \frac{p_i}{\mu_i}) \quad (5)$$

- **JACCARDINDEX** – индексът на Джакард се изразява в намирането на относителния брой (процента) на общите елементи спрямо броя на всички, без дублиранятия, в два вектора. 0

Векторът  $Q = (q_1, q_2, \dots, q_n)$  представлява изображението заявка, а векторът  $P = (p_1, p_2, \dots, p_n)$  – произволно изображение.







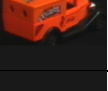


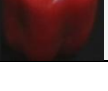
Векторът  $\mu = (\mu_1, \mu_2, \dots, \mu_n)$  се получава като:  $\mu = \frac{p+q}{2} \cdot 0$









Към вече споменатите мерки за сходство ние добавихме още една – събиране на разликите от нашия цетови модел с 64 цвята между двете изображения позиционно. После резултатите са съпоставени с резултатите от останалите мерки за сходство.

### 3. Резултати

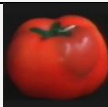
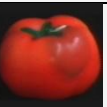
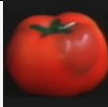
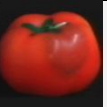
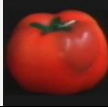

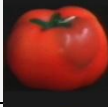
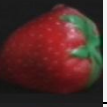
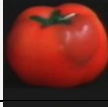

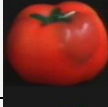

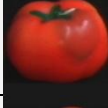
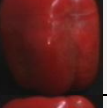
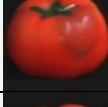
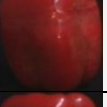
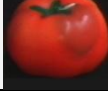
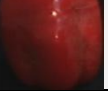
От двете множества с изображения се взема произволно по едно изображение (заявка), което е сравнено с всички останали. Важно е да се отчете, че всички резултати са мащабирани (нормализирани) и показани в процентно съотношение. В таблица 1 и таблица 2 са показани сравняваните изображения и резултатите за степента на подобие, според различните мерки за сходство, като от ляво на дясно са разположени заявката изображение, отделните изображения в множеството, които се сравняват със заявката, следвани от отделните мерки за сходство, които са дадени като стойност за сходството между двете изображения в първите две колони в проценти и с индекс в скоби, показващ подредбата по степен на подобие.

**Таблица 1.** Резултати от сравнението на изображения с различни обекти и техните проценти

Query Image	Search Image	Jaccard	Euclidean Distance	Cityblock Distance	Chi-square Dissimilarity	Kullback-Diebler Divergence	Jeffrey Divergence	My Array Differences
		(1) 100.0000	(1) 100.0000	(1) 100.0000	(1) 100.0000	(1) 100.0000	(1) 100.0000	(1) 100.0000
		(2) 74.1211	(6) 75.3952	(2) 89.0470	(4) 91.8441	(150) 50.8475	(3) 92.2930	(4) 99.5626
		(3) 69.4336	(54) 68.5930	(27) 84.1797	(35) 86.7444	(193) 30.7825	(40) 87.2650	(5) 99.5385
		(96) 47.2656	(150) 59.5795	(129) 75.0667	(139) 78.0120	(126) 58.8525	(144) 78.5414	(81) 98.8977
		(125) 44.2383	(31) 70.6438	(74) 79.4054	(63) 84.3367	(90) 68.9959	(62) 85.4059	(84) 98.8865

			(126) 43.9453	(172) 56.9767	(162) 71.6967	(177) 74.5749	(148) 51.8022	(178) 75.1894	(113) 98.7637
			(214) 21.4844	(215) 43.4418	(215) 54.2814	(216) 58.5133	(108) 63.6046	(216) 59.5224	(143) 98.5257
			(215) 21.0938	(213) 43.9430	(214) 54.3728	(215) 58.7269	(128) 58.6051	(215) 59.8004	(200) 98.0886
			(216) 17.5781	(214) 43.5530	(216) 54.1341	(214) 59.3517	(78) 73.1851	(214) 60.6498	(186) 98.2184

**Таблица 2.** Резултати от сравнението на изображения с еднакви обекти и техните проценти

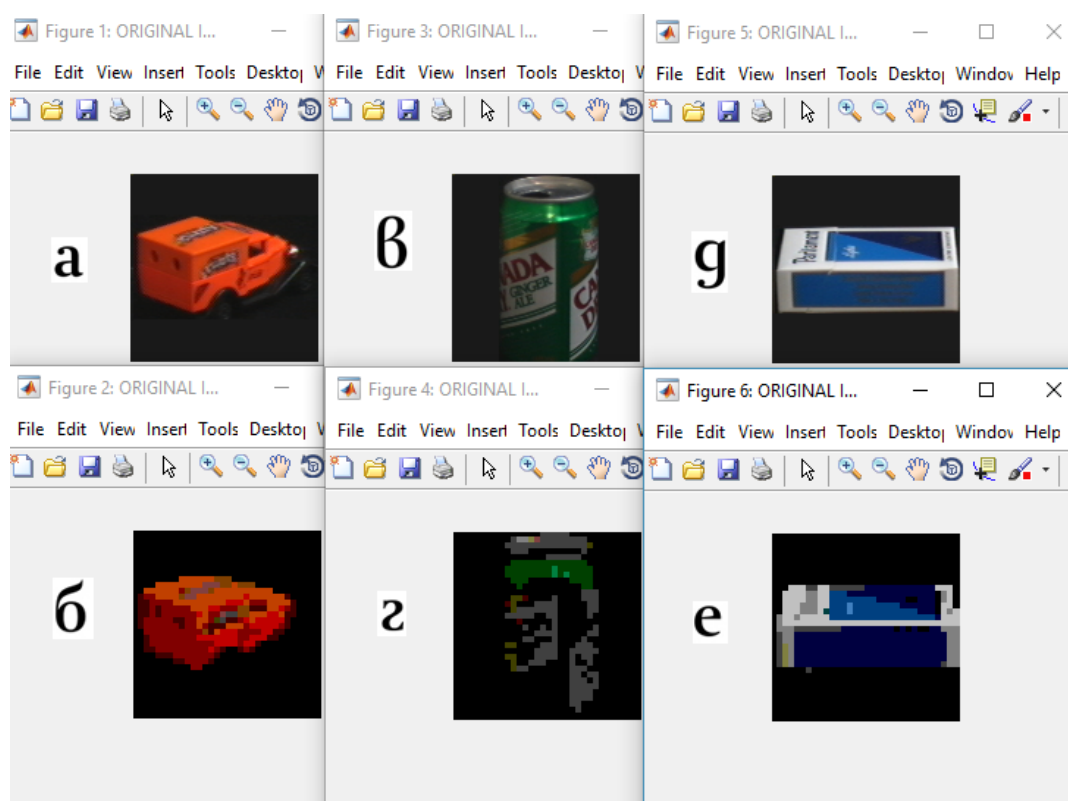
Query Image	Search Image	Jaccard	Euclidean Distance	Cityblock Distance	Chi-square Dissimilarity	Kullback-Divergence	Jeffrey Divergence	My Array Differences
		(1) 100.0000	(1) 100.0000	(1) 100.0000	(1) 100.0000	(1) 100.0000	(1) 100.0000	(1) 100.0000
		(2) 88.2813	(2) 90.7220	(2) 97.1292	(2) 98.7747	(3) 95.0487	(2) 98.9282	(2) 99.9081
		(3) 87.2070	(3) 89.8507	(3) 96.7587	(3) 98.7195	(4) 92.8743	(3) 98.8951	(16) 99.8193
		(96) 47.7539	(109) 74.3743	(95) 83.4248	(93) 90.7231	(152) 30.6821	(91) 91.8402	(176) 98.9849
		(125) 44.7266	(108) 74.3926	(92) 83.5751	(94) 90.7209	(101) 43.8490	(92) 91.8288	(204) 98.7481
		(126) 44.5313	(91) 75.3269	(108) 83.1272	(102) 90.1355	(78) 57.9260	(102) 91.1148	(88) 99.3395
		(214)33.6 914	(200) 69.8089	(213) 76.9345	(209) 85.0049	(160) 28.3768	(208) 86.4060	(134) 99.1386
		(215) 33.3008	(208) 69.2419	(216) 76.4307	(213) 84.6223	(162) 28.1298	(212) 86.0811	(157) 99.0772
		(216) 32.8125	(207) 69.3537	(215) 76.4741	(211) 84.6815	(157) 29.6630	(211) 86.1341	(156) 99.0824

#### 4. Изводи

От анализа на резултатите на направените изследвания се вижда, че при различните мерки за сходство има разлика в подредбата на изображенията и техните проценти. Всяка мярка за сходство по свой начин представя резултатите от анализираното множество, но, въпреки разминаванията, подредбата е почти една и съща в повечето метрики. Въпреки това се откроява „KULLBACK-LIEBLER DIVERGENCE“ мярката. При нея се наблюдава голямо несъответствие с останалите използвани мерки. Това се дължи на факта, че тази мерка е логаритмична и е необходимо да се използва абсолютната стойност на резултата, преди да се преобразува в проценти. Поради тази причина имаме резултати, които се появяват като много подобни, но това се дължи на факта, че всъщност абсолютната стойност на този резултат е подобен, но не и самото изображение, което в крайна сметка се оказва тотално различно като цветове.

Друг важен фактор, който може да отчетем от направеното сравнение, е, че мерките за сходство дават много по-точни резултати при множеството с почти еднакви изображения (съставено от 3 обекта), отколкото при множеството с различни изображения (съставено от 100 обекта). За това вероятно има принос и фонът. Защото той дава голям процент от стойността на подобие. Ако се игнорира фонът, вероятно нещата ще се подобрят и за множеството от произволни изображения, с всякакви обекти. Както се вижда от таблица 2, повечето от мерките успешно разпознават дори отделните обекти и успяват да ги подредят последователно, т. е. да разпознаят, че заявката напълно съответства на най-близките резултати като обект и да го индексира като най-близък резултат на едно от първите места, а след това да подреди останалите два обекта отново обособени един от друг.







От направеното сравнение също се наблюдава, че част от метриките дават изключително близки стойности на обекти, които, въпреки общата си прилика, са ориентирани различно в пространството. Това може да се види на фигура 3.



Фиг. 3. Генериране на нови изображения с редуцирани цветове с 64 различни цветови разновидности

От фигура 3 се забелязва, че въпреки почти хоризонталната ориентация на обекта в изображението-заявка (3а и 3б) и вертикалната ориентация на обекта в изображението, с което се сравнява (3в и 3г), процентите на подобие между 3а и 3в, и 3а и 3д са много близки. Причината за това е, че при конвертиране до 64 цвята се наблюдава изменение на някои от участъците в обекта от 3в/3г и от тъмно зелени те стават черни. Поради това, дори при перпендикулярна ориентацията на обекта и позиционно сравнение, ще се получат близки резултати, защото въпреки малките разлики в цветовете на двете изображения, те попадат в един диапазон на черния цвят при конвертиране. Сравнението с мерките за сходство може да се види в Таблица 3.

**Таблица 3.** Резултати от сравнение на изображения с различно пространствено разположение

Query Image	Search Image	Jaccard	Euclidean Distance	Cityblock Distance	Chi-square Dissimilarity	Kullback-liebler Divergence	Jeffrey Divergence	My Array Differences
		(1) 100.0000	(1) 100.0000	(1) 100.0000	(1) 100.0000	(1) 100.0000	(1) 100.0000	(1) 100.0000
		(33) 57.5195	(85) 65.6704	(53) 81.2764	(54) 84.9230	(176) 39.5730	(54) 86.0042	(122) 98.7311
		(34) 56.2500	(69) 67.2313	(59) 80.7788	(80) 82.7506	(215) 3.2640	(80) 83.4526	(39) 99.1042

### Литература

- [1]. Collins J., K. Okada, A Comparative Study of Similarity Measures for Content-Based Medical Image Retrieval, San Francisco State University, 1600 Holloway Avenue, San Francisco, CA 94132, USA, [https://bidal.sfsu.edu/~kazokada/research/okada\\_clef12\\_mcbir.pdf](https://bidal.sfsu.edu/~kazokada/research/okada_clef12_mcbir.pdf), с. 2–7.
- [2]. Jaccard, Paul (1901), "Étude comparative de la distribution florale dans une portion des Alpes et des Jura", Bulletin de la Société Vaudoise des Sciences Naturelles, 37: 547–579.
- [3]. Nene, S. A., S. K. Nayar and H. Murase, Columbia University Image Library (COIL-100), Technical Report CUCS-006-96, February 1996, <http://www1.cs.columbia.edu/CAVE/software/softlib/coil-100.php>, с. 1.
- [4]. Valova, I., B. Rachev, Image organization, querying and retrieval by color distribution features, International Conference CompSysTech'2002, Sofia, Bulgaria, 20-21 June 2002.
- [5]. Vassilakopoulos M., Corral A., Rachev B., Valova I. & Stoeva M., IMAGE DATABASE INDEXING TECHNIQUES, Handbook of Research on Geoinformatics, 20-27.

### За контакти:

маг. инж. Мирослав Б. Маринов  
катедра „Компютърни системи и технологии“  
Русенски университет „Ангел Кънчев“  
E-mail: mbmarinov@uni-ruse.bg

# НОВИ НЕВРОННИ МОДЕЛИ НА ЛРВНМ БАЗИРАНИ НА RELU АКТИВИРАЩИ ФУНКЦИИ

Николай Т. Дуков, Тодор Д. Ганчев

**Резюме:** Невроните в рекурентния слой на локално рекурентната вероятностна невронна мрежа традиционно използват сигмоидална активираща функция. Основен недостатък на тази активираща функция е значителната ѝ изчислителна сложност и необходимостта от трудни за имплементиране в FPGA нелинейни функции. В настоящата разработка се предлага нова активиращата функция от семейството на ReLU, наречена SatReLU, която комбинира положителните характеристики на Leaky ReLU и Clipped ReLU. Предимствата, произтичащи от използването на предложените нови невронни модели, са изследвани с помощта на експериментална постановка за разпознаване на негативни емоционални състояния от физиологични сигнали, реализирана с помощта на базата данни DEAP.

**Ключови думи:** невронни мрежи, EEG, LRPNN, ReLU, активираща функция, DEAP.

## Novel ReLU-based neuron models for the LRPNN

Nikolay T. Dukov, Todor D. Ganchev

**Abstract:** The locally recurrent probabilistic neural network (LRPNN) traditionally uses a sigmoidal activation function in the recurrent layer neurons. A disadvantage of this function is the computational complexity and challenging implementation in FPGA designs. In the current study we investigate alternatives based on the ReLU activation function and its modifications. We propose a new member of the ReLU family, referred to as SatReLU, which combines the advantages of Leaky ReLU and Clipped ReLU. The experimental validation on the negative emotion recognition task, carried out on the DEAP database, confirms the advantages of the proposed neuron models.

**Keywords:** neural network, EEG, LRPNN, ReLU, activation function, DEAP.

## 1. Въведение

Локално рекурентната вероятностна невронна мрежа (ЛРВНМ) открива и моделира времеви зависимости в последователности от данни, което ѝ дава предимства в приложения, свързани с откриване на събития и разпознаване на поведения, като например разпознаване на диктори или емоционални състояния от гласови данни или EEG сигнали [1], [2].

Същевременно, от гледна точка на възможностите за апаратна реализация на ЛРВНМ, с цел подобряване на бързодействието ѝ, като основен недостатък на традиционната архитектура се отчита използването на сложни нелинейни зависимости в активиращите функции на невроните от образцовия слой и рекурентния слой. Апаратната реализация на тези активиращи функции се явява основно предизвикателство при имплементации с помощта на програмируеми цифрови платформи (FPGA), като най-често се търсят компромиси между размера на ресурсите, използвани за реализация на нелинейните активиращи функции и точността при класификация. Логично решение на проблема е замяната на някои от нелинейните активиращи функции с други, които са по-удобни за апаратна реализация. В тази връзка, в последните години набират популярност нелинейни предавателни функции от семейството ReLU (Rectified Linear Units), които имат редица предимства спрямо традиционно използваните. Тези предимства са показани в приложения, свързани с гласово разпознаване [3], [4], [5] и други. Тук подобренията не се изчерпват единствено до процентно намаляване на грешката при разпознаване, а също така се намалява времето за обучение на класификатора. Предимствата на активиращите

функции от семейството на ReLU, в т.ч. нейните модификации Leaky ReLU и Clipped ReLU, произтичат от начина, по който е дефинирана нелинейността, както и от ефективното им изчисление, особено в сравнение с традиционната сигмоидална активираща функция.

В Секция 2 на настоящата разработка се предлага описание на традиционната архитектура на ЛРВНМ и модификациите, свързани с въвеждане на новата SatReLU активираща функция, която комбинира положителните характеристики на Leaky ReLU и Clipped ReLU. Предимствата, произтичащи от използването на предложените нови невронни модели, са изследвани с помощта на експериментална постановка за разпознаване на негативни емоционални състояния от физиологични сигнали, представена в Секция 3, която е реализирана с помощта на базата данни DEAP. В Секция 4 е представен анализ на експерименталните резултати.

## 2. Изложение

### 2.1. Архитектура на Локално Рекурентната Вероятностна Невронна Мрежа

Локално рекурентната вероятностна невронна мрежа [2] надгражда вероятностната невронна мрежа [6] чрез внедряване на допълнителен скрит слой, състоящ се от напълно свързани рекурентни неврони, наречен рекурентен слой. Така формираната хибридна архитектура дава възможност за долавяне и моделиране на времеви корелации и взаимовръзки в последователности от входни данни, което носи определени предимства в задачи, свързани с класификация на събития и поведения [2], [7], [8].

Архитектурата на ЛРВНМ е представена на фигура 1 [8]. Рекурентният слой се явява четвърти, след входния слой, слоя с образци и сумиращия слой. Пети, след рекурентния слой, е изходният слой, който реализира правилото на Бейс за класификация с минимален риск. Както и при традиционната ВНМ, за изчисление на вероятността за всеки клас са използвани прозорци на Парзен, реализирани чрез Гаусови функции [2]. Изходът на сумиращия слой  $f_i(x_p)$  (1) се явява вход за рекурентния слой. По принцип рекурентният слой може да се окачестви като филтър с безкрайна импулсна характеристика. Неговата роля е да изглажда стойностите, генерирани за отделните класове, като важна особеност е, че това се извършва и за една или повече предишни стойности.

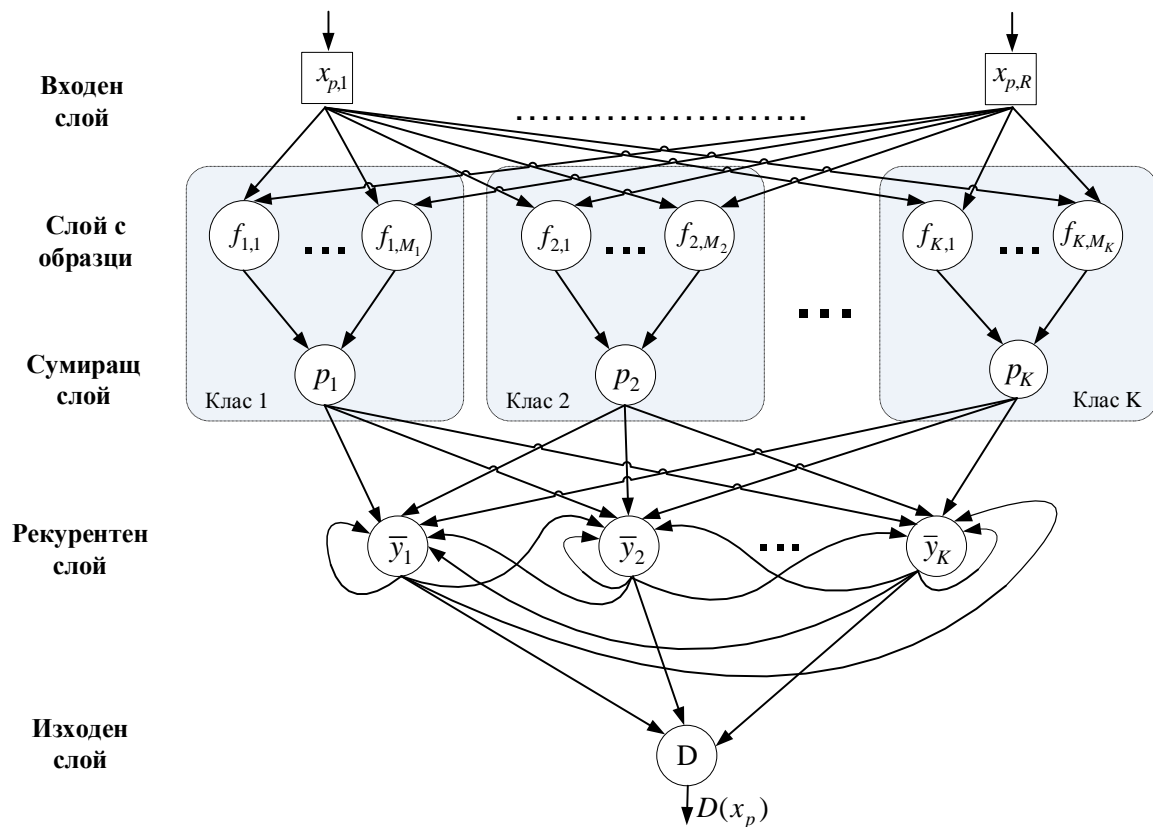
При класификацията на данни чрез ЛРВНМ, функцията на разпределение на вероятността се изчислява както [2]:

$$f_i(x_p) = \frac{1}{(2\pi)^{d/2} \sigma_i^d} \cdot \frac{1}{M_i} \sum_{j=1}^{M_i} \exp\left(-\frac{1}{2\sigma_i^2} (x_p - x_{ij})^T (x_p - x_{ij})\right), \quad i=1, \dots, K. \quad (1)$$

Съответно с  $x_{ij}$  е отбелязан  $j$ -тият обучаващ вектор от клас  $k_i$ , а с  $x_p$  - входният тестов вектор. Тук  $d$  е размерността на вектора, а  $M_i$  е броят тренировъчни вектори в клас  $k_i$ . Предполага се, че всеки обучаващ вектор  $x_{ij}$  е център на ядрото на класификатора. С  $\exp$  е означена експоненциалната функция, а индексът  $T$  оказва транспониране. Стандартното отклонение  $\sigma$  играе ролята на изглаждащ фактор, „омекотяващ“ повърхността, дефинирана от множеството Гаусови функции.

Рекурентните неврони получат входни стойности от предишния сумиращ слой, които се сумират със свои собствени предишни състояния, както и предишни стойности от изхода на неврони на други класове. Изходът от сумиращите неврони в локално рекурентния слой се изчислява както следва [8]:





Фиг. 1. Архитектура на ЛРВНМ [8]

$$y_i(x_p) = \left[ b_{i,i} f_i(x_p) - \sum_{\substack{k=1 \\ i \neq k}}^K b_{i,k} f_k(x_p) \right] + \sum_{t=1}^N \left[ a_{i,i,t} \bar{y}_i(x_{p-t}) - \sum_{\substack{k=1 \\ i \neq k}}^K a_{i,k,t} \bar{y}_k(x_{p-t}) \right], \quad (2)$$

$i = 1, \dots, K$

където  $f_i(x_p)$  е функцията на разпределение на вероятността за всеки определен клас  $k_i$ ,  $x_p$  – входния вектор,  $K$  – броя класове,  $N$  – рекурентната дълбочина,  $y_i(x_{p-t})$  е нормализираната предишна изходна стойност за клас  $k_i$ , която е била задържана с  $t$  време,  $a_{i,j,t}$  и  $b_{i,j}$  от своя страна са тегловни коефициенти. Изходът на сумиращия неврон  $y_i(x_p)$  в рекурентния слой се подлага на нормализация, която запазва вероятностната интерпретация на изхода на рекурентния слой. С  $sgm$  е означена сигмоидалната активираща функция [8]:

$$\bar{y}_i(x_p) = \frac{sgm(y_i(x_p))}{\sum_{j=1}^K sgm(y_j(x_p))}, \quad i = 1, 2, \dots, K \quad (3)$$

За активираща функция в рекурентния слой на ЛРВНМ [2] се използва сигмоидалната активираща функция (4), която приема реални числа и връща като резултат положителни числа в интервала  $[0, 1]$ :

$$f(x) = \frac{1}{(1 + \exp^{-x})} \quad (4)$$

Сигмоидалната предавателна функция може да се определи като функция с нелинейно активиране. Нивото на възбуждане се изменя в най-голяма степен при малки входни стойности, докато при големи положителни или отрицателни стойности се наблюдава плавно

приближаване към долната или горна граница, т.е. значителни изменения на числените стойности на входа предизвикват минимални изменения в изхода.

В изходния слой, известен и като конкурентен слой, се прилага правилото на Бейс за класификация с минимален риск за определяне на класа, към който принадлежи даден входен вектор [8]:

$$D(x_p) = \arg \max_i \{h_i c_i \bar{y}_i(x_p)\}, i = 1, \dots, K \quad (5)$$

### 2.1.1. Нови активиращи функции за ЛРВНМ

#### *ЛРВНМ със SatReLU активираща функция*

В настоящата разработка предлагаме нова активираща функция за невроните от рекурентния слой на ЛРВНМ, която в значителна степен намалява изчислителната сложност и елиминира необходимостта от трудната за имплементиране в FPGA сигмоидална активираща функция. Предложената активираща функция, SatReLU, принадлежи към семейството на ReLU функциите и комбинира положителните качества на Leaky ReLU и на Clipped ReLU:

$$f(x) = \begin{cases} 0.001 \cdot const, & x < 0 \\ x, & 0 \leq x < const \\ const, & x \geq const \end{cases} \quad (9)$$

където *const* е максимална стойност на насищане, над която всички стойности биват приравнени на *const*. Стойността на *const* също се използва за дефиниране на праг за отрицателните стойности на входа, който обаче е значително по-малък по стойност.

В допълнение, за нуждите на сравнителен експериментален анализ, разглеждаме използването и на други ReLU функции, като възможни активиращи функции на невроните от рекурентния слой.

#### *ЛРВНМ със стандартна ReLU активираща функция*

Накратко реакцията на ReLU активационната функция при отрицателен вход е нула на изхода. Положителната стойност на входа на ReLU активиращата функция се повтаря от изхода [9]:

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (6)$$

#### *ЛРВНМ с Leaky ReLU активираща функция*

Leaky ReLU [10] е модифициран вариант на ReLU:

$$f(x) = \begin{cases} x, & x \geq 0 \\ const \cdot x, & x < 0 \end{cases} \quad (7)$$

където *const* е фиксирана скаларна стойност, която пропуска отрицателни стойности с цел избягване на насищането в долната граница. Обикновено параметърът *const* е с малки стойности, от порядъка на 0.01.

#### *ЛРВНМ с Clipped ReLU активираща функция*

Clipped ReLU [5] е модификация на ReLU:

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x < const \\ const, & x \geq const \end{cases} \quad (8)$$

където в този случай  $const$  е максимална стойност, над която всички стойности биват приравнени на  $const$ .

Изходът на сумиращия неврон от рекурентния слой е подложен на нормализация за запазване на вероятностната интерпретация при използване на активиращи функции от фамилията на ReLU:

$$\bar{y}_i(x_p) = \frac{|rl(y_i(x_p))|}{\sum_{j=1}^K |rl(y_j(x_p))|}, \quad i=1,2,\dots,K \quad (9)$$

където с  $rl$  е означена активираща функция от семейството на ReLU.

### 2.1.2. Обучение на Локално Рекурентната Вероятностна Невронна Мрежа

Обучението на ЛРВНМ протича в три стъпки [8]. Първата стъпка е аналогична с тази при обучението на ВНМ, която създава топологията на мрежата. В първия скрит слой се формират тегловни вектори, които са равни по брой на тренировъчните вектори. Изходите на даден клас от първия скрит слой се свързват с някой от сумиращите неврони във втория скрит слой. Броят на сумиращите неврони от своя страна е равен на броя класове.

Втората стъпка по обучението се състои в изчислението на изглаждащия параметър  $\sigma_i$  за всеки един от класовете или общо за всички класове [11], [12], [8].

Последната, трета стъпка е свързана с изчислението на тегловните коефициенти на локално рекурентния слой, което включва минимизиране на функция на грешката [2], [7], [8]. За тази стъпка е необходимо използването на тренировъчните данни от стъпка едно. Съответно за изчисление на тегловните коефициенти е използван алгоритъмът диференциална еволюция, който постига добри резултати при приемливо изчислително време [8]:

$$E(w) = \sum_{i=1}^K m_i (1 - P(k_i | X_{k_i})) P(k_i) + \frac{1}{K-1} \sum_{i=1}^K \sum_{\substack{j=1 \\ j \neq i}}^K m_j P(k_j | X_{k_i}) P(k_j), \quad (10)$$

където  $i = 1, 2, \dots, K$  е индексът на класа, а  $w$  е векторът с тегловни коефициенти на рекурентния слой, подлежащи на настройка.

## 3. Експериментална постановка

За нуждите на експерименталното изследване са дефинирани пет различни невронни модела на ЛРВНМ, които използват различни активиращи функции в рекурентния слой: (i) Sigmoid (сигмоидална), (ii) ReLU, (iii) Leaky ReLU, (iv) Clipped ReLU и (v) SatReLU. За моделите с Leaky ReLU, Clipped ReLU и SatReLU са разгледани по десет варианта на параметъра  $const$ . Съответно за Leaky ReLU това са стойностите от 0.01 до 0.1 със стъпка 0.01, а за Clipped ReLU и SatReLU от 10 до 100 със стъпка 10.

За обучение и проверка поведението на различните модели на неврони в рекурентния слой е използвана базата данни DEAP [13]. Опитната постановка е насочена към разпознаване на *негативни емоционални състояния* и следва експерименталният протокол, описан в [14], [15]. Накратко, използвани са ЕЕГ записи от десет човека, като

класификацията се базира на изчислената нормализирана кратковременна енергия от 32-канално ЕЕГ. Детекторът разпознава два класа: (а) *негативни* емоционални състояния, и (б) обобщен клас *други*, включващ неутрални и позитивни емоционални състояния. ЕЕГ записите са разделени в три групи: обучаващи данни – 20%; развойни данни – 20%; тестови данни – 60%.

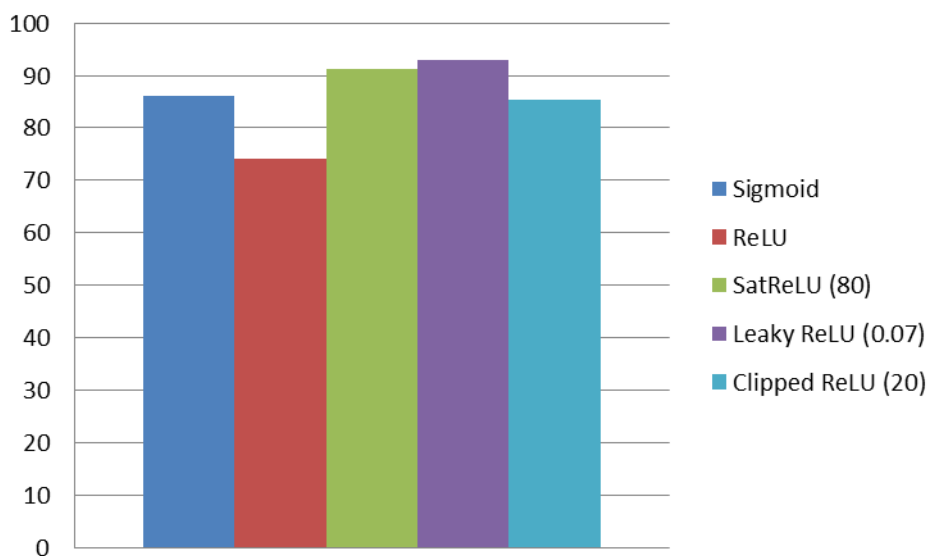
Обучаващите данни са използвани за създаване на топологията на невронна мрежа и формиране на образцовия слой на ЛРВНМ. Изчислението на изглаждащия фактор  $\sigma$ , както и изчислението на тегловните коефициенти на рекурентния слой е извършено чрез използването на развойните данни. При изчислението на изглаждащия фактор  $\sigma$  се избира малка начална стойност, която плавно се увеличавана до намиране на оптималната такава.

За изчисление на тегловните коефициенти в рекурентния слой е използван оптимизиращия алгоритъм DE. Големината на популацията е зададена като 20 пъти размерността на оптимизирания вектор от тегловни коефициенти. Оптимизацията се извършва без ограничения в разпространението на популацията в пространството на решенията. Итерациите на оптимизационния алгоритъм се прекратяват, когато целевата функция (10) достигне нула или при достигане на максималния брой итерации - 2000. Параметрите на DE, наречени коефициенти на рекомбинация и мутация, са зададени равни на 0.8. Началните условия на оптимизационния алгоритъм DE са зададени след проведени обширни експерименти и на база на предходни знания, като е използван операторът DE/rand-to-best/1/bin, който показва най-добри резултати за конкретната оптимизационна задача.

За всеки от петте невронни модела и съответните стойности на параметъра *const* са извършени по десет независими обучения на рекурентния слой.

#### 4. Резултати

Обобщените резултати от експерименталните изследвания са показани на фигура 2. От гледна точка на критерия *точност при класификация* се наблюдава преимущество на две от ReLU модификациите спрямо модела на рекурентен неврон със сигмоидална активираща функция. Докато при сигмоидалната функция средната стойност за *точността на класификация* е 86.2%, то при Leaky ReLU е 92.8% (за *const* = 0.07), а при SatReLU е 91.4% (за *const* = 80). При същите условия стандартната ReLU функция постига максимален среден резултат 74.1%, а Clipped ReLU 85.4% при *const* = 20.



Фиг. 2. Обобщени резултати от проведените експерименти.

Невронният модел със стандартната ReLU функция показва най-ниските резултати от изследваните модели, което в голяма степен се дължи в непропускането на постъпващите отрицателни стойности от активиращата функция и в известна степен на липсата на ограничение за положителните. Това следва от резултатите, наблюдавани за Leaky ReLU и новата SatReLU, за които са наблюдавани най-високата точност – вземайки под внимание дефиницията им с (7) и (9) отчитаме, че и двете активиращи функции пропускат отрицателни стойности. За разлика от Leaky ReLU, която пропуска отрицателни стойности, като ги умножава с малък коефициент, то при SatReLU отрицателните стойности се заместват с малка фиксирана положителна стойност. Това по същество е сходно с начина, по който стандартната ReLU активираща функция и Clipped ReLU активиращата функция предават отрицателните входни стойности. За разлика от SatReLU обаче, отрицателните стойности при тях водят до изходна стойност нула. Също така при SatReLU има възможност за настройка на стойността, която ще бъде получена при отрицателен вход. Изчислението на тази стойност е обвързано със стойността, определена за праг, при постъпването на положителни стойности (9). От една страна това може да се разглежда като недостатък, но същевременно по този начин се намалява броят на необходимите параметри за настройка. В заключение можем да обобщим, че при подходящ избор на стойност за параметъра *const*, невронният модел с функцията SatReLU превъзхожда моделите с ReLU и Clipped ReLU.

За разлика от разгледаните ReLU активиращи функции, Sigmoid функцията променя динамичния диапазон на постъпващите данни (5), което е и голямото предимство на тази функция. Постъпващи входни данни в интервала  $[-\infty, +\infty]$  биват преобразувани до изходни в диапазона  $[0, 1]$ . За съжаление, поради ограничения в броя битове при представяне на стойностите на Sigmoid (например при имплементация с целочислена аритметика), изходът на функцията за определени входни стойности достига границите 0 или 1. Това от своя страна е съпроводено със загуба на информация, което се потвърждава и от експерименталните резултати (фигура 2).

Това важи с още по-голяма сила при хардуерна имплементация в FPGA, където често данните в приложения с невронни мрежи се представят във формат с фиксирана десетична запетая, с цел редуциране на необходимите апаратни ресурси. В допълнение, пресмятането на Sigmoid чрез ред на Маклорен е изчислително интензивно, което прави обучението на невронна мрежа по-бавно, отколкото ако се използват активиращи функции от семейството на ReLU.

## Заклучение

Използването на нови активиращи функции за невроните от рекурентния слой на ЛРВНМ носи преимущества, състоящи се в подобрена точност при класификация, намалена сложност на изчисленията и подобрена скорост на обучение на рекурентния слой. Подобрена точност при класификация е постигната за Leaky ReLU с *const* = 0.07 и предложената SatReLU с *const* = 80. Активиращата функция SatReLU също демонстрира по-добри показатели от гледна точка на предвидимостта спрямо останалите изследвани функции, което с оглед на малката разлика ( $\approx 1\%$ ) с Leaky ReLU в постигнатата точност я прави добра нейна алтернатива. Преимущество, валидно както за ReLU, така и за разгледаните нейни модификации, е скоростта на обучение, което е приблизително 3 пъти по-бързо отколкото при сигмоидалната функция. Тези преимущества показват, че предложените модификации на ЛРВНМ са особено подходящи при апаратна реализация с помощта на цифрови платформи от рода на FPGA.

## Литература

- [1]. N. Dukov, T. Ganchev, D. Kovachev, „FPGA implementation of the Locally Recurrent Probabilistic Neural Network“. *Advances in Intelligent Systems and Computing*, vol.679, pp. 419-428, 2018.
- [2]. T. Ganchev, D.K. Tasoulis, M.N. Vrahatis, N. Fakotakis, „Locally recurrent probabilistic neural network for text-independent speaker verification,“ *Proc. of the 8th European Conference on Speech Communication and Technology, EUROSPEECH 2003*, vol.3, pp. 1673-1676, 2003.
- [3]. P. Ramachandran, Z. Barret, V.L. Quoc, „Searching for Activation Functions,“ *arXiv:1710.05941*, 2017.
- [4]. A.F. Agarap, „Deep Learning using Rectified Linear Units (ReLU),“ *arXiv preprint arXiv:1803.08375*, 2018.
- [5]. C. Case, J. Casper, B. Catanzaro, A. Coates, G. Diamos, E. Elsen, A. Y. Hannun, A. Y. Ng, R. Prenger, S. Satheesh, S. Sengupta, „Deep Speech: Scaling up end-to-end speech recognition,“ *arXiv preprint arXiv:1412.5567*, 2014.
- [6]. D. Specht, „Probabilistic Neural Networks,“ *Neural Networks*, vol.3, pp.109-118, 1990.
- [7]. T. Ganchev, D.K. Tasoulis, M.N. Vrahatis, N. Fakotakis, „Locally recurrent probabilistic neural networks with application to speaker verification,“ *GESTS Int. Trans. Speech Sci. Eng.*, vol.1, no. 2, pp.1-13, 2004.
- [8]. T. Ganchev, „Enhanced training for the locally recurrent probabilistic neural networks,“ *International Journal on Artificial Intelligence Tools*, vol.18, no.6, pp.853-881, 2009.
- [9]. V. Nair, G.E. Hinton, „Rectified linear units improve restricted Boltzmann machines,“ *Proc. of the 27th international conference on machine learning (ICML-10)*, pp.807-814, 2010.
- [10]. A.L. Maas, A.Y. Hannun, A.Y. Ng, „Rectifier nonlinearities improve neural network acoustic models,“ *In Proc. ICML*, vol.30, no.1, 2013.
- [11]. J.B. Cain, „Improved probabilistic neural network and its performance relative to other models,“ *Proceedings SPIE, Applications of Artificial Neural Networks*, vol.1294, pp.354-365, 1990.
- [12]. D. F. Specht, „Enhancements to Probabilistic Neural Networks,“ *Proc of IEEE International Joint Conference on Neural Networks*, том 1, pp. 761-768, 1992.
- [13]. S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras, „DEAP: A Database for Emotion Analysis using Physiological Signals,“ *IEEE Transactions on Affective Computing*, vol.3, no.1, pp.18-31, 2011.
- [14]. F. Feradov, T. Ganchev, „Detection of negative emotional states from Electroencephalographic (EEG) signals,“ *Annual Journal of Electronics*, vol.8, pp.66-69, 2014.
- [15]. N. Dukov, T. Ganchev. „Empirical study on ReLU based neuron models of the LRPNN“. *Biomedical Data Acquisition and Applications Workshop, Varna, Bulgaria*, pp.24-27, 2017.

**За контакти:**

ас. инж. Николай Т. Дуков  
катедра „Компютърни науки и технологии“  
Технически университет-Варна  
E-mail: nikolay.dukov@tu-varna.bg

# ОБРАБОТКА НА ТРИМЕРНИ ОБЕКТИ СЕГМЕНТИРАНИ ОТ МЕДИЦИНСКИ ИЗОБРАЖЕНИЯ

Гергана В. Спасова, Николай Т. Дуков, Кристина С. Близнакова

**Резюме:** В тази разработка е представен метод за получаване на различни туморни образувания сегментирани от реални медицински изображения. Използвани са пациентски изображения получени от 3D маммография. Приложен е метод за полуавтоматично извличане на туморни образувания от тези изображения и представянето им като вокселна 3D матрица. Използвани са набор от морфологични операции върху тримерните туморни модели, като в резултат на това са получени множество разнообразни нови гърдни лезии.

**Ключови думи:** обработка на изображения, тримерно изображение, лезии, сегментиране

## Processing of three-dimensional objects segmented from patient medical images

Gergana V. Spasova, Nikolay T. Dukov, Kristina S. Bliznakova

**Abstract:** In this paper, a method for generation of various tumor formations, based on segmented abnormalities from patient medical images is presented. Patient images are derived from three-dimensional mammography images. Initially, a method is used to semi-automatically extract tumor formations from these images and represent them as voxel matrix. A set of morphological operations was applied on these voxel tumor matrixes which resulted in a variety of new breast lesions.

**Keywords:** image processing, 3D image, lesions, segmentation

### 1. Увод

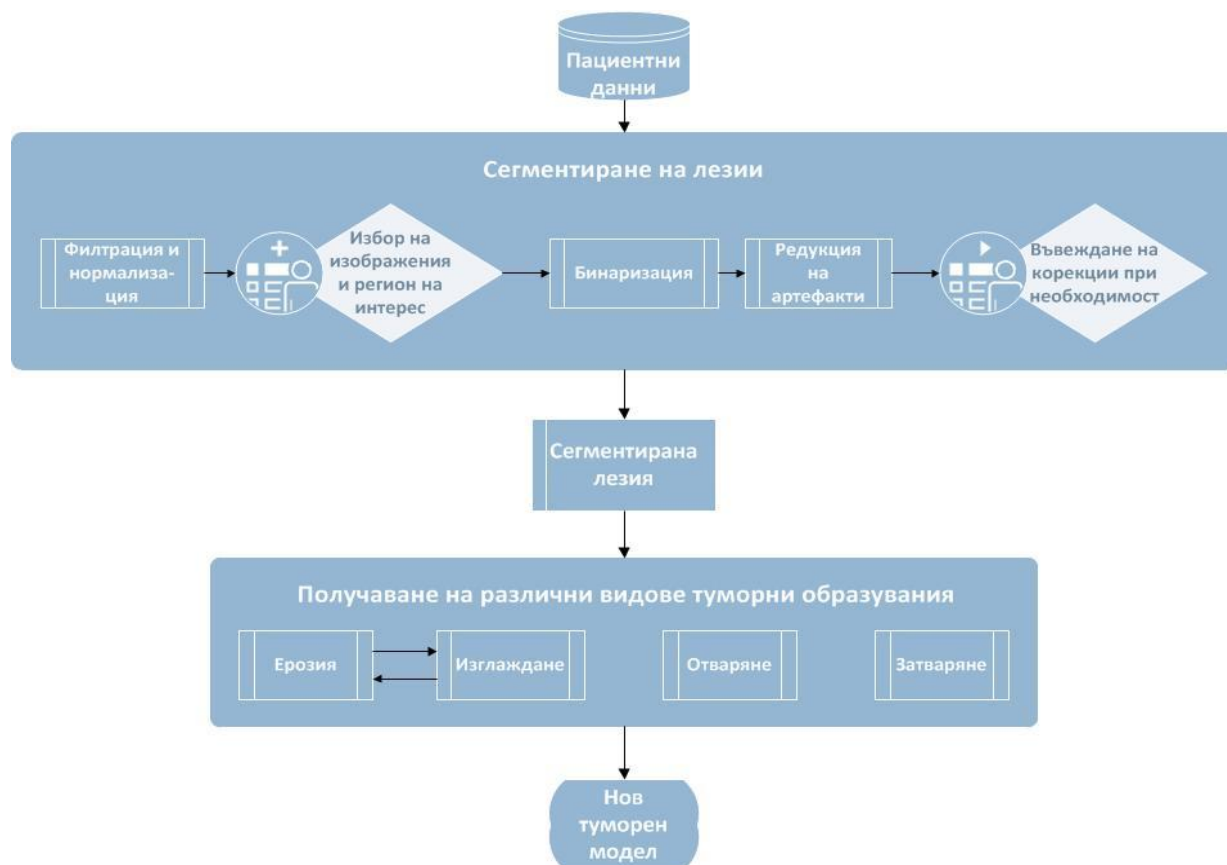
Ракът на гърдата е най-широко разпространения вид рак сред жените. В България той е причина за смъртта на всяка една от осем жени [1]. Това налага необходимостта от разработката на нови технологии, свързани с неговото навременно откриване и диагностициране.

Разработването на нови техники за скрининг и диагностика на женската гърда изискват реалистични туморни образувания, които да бъдат внедрени в компютърни или физически модели на гърда. Проблемът е в това, че за разработването на такива техники са нужни много и различни видове туморни образувания, до които изследователите нямат достъп поради липсата на информация.

Този доклад представя метод за генериране на туморни образувания на базата на такива, сегментирани от пациентски данни. Чрез прилагане на набор от морфологични операции е възможно получаване на различни и реалистични видове туморни образувания за целите на разработването на нови техники за скрининг и диагностика на рак на гърдата и за целите на системите за автоматизирано и компютърно проектиране (Computer Aided Design (CAD)).

### 2. Изложение

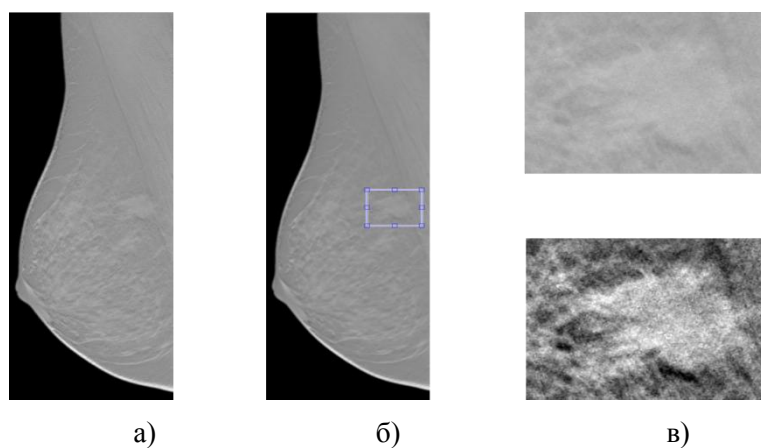
Процесът на генериране на туморни образувания включва няколко етапа, които са показани на фигура 1.



**Фиг. 1.** Диаграма на процеса на генериране на туморни образувания

## 2.1. Пациентски данни

Използваните пациентски данни са предоставени от две болници – Университетска болница Льовен, Льовен, Белгия и Университетска болница „Александровска“, София, България. Тези данни са под формата на томографски мамографски изображения, получени съответно от машините Siemens Mammomat Inspiration и Giotto Tomo IMS. Данните са анонимизирани, като единствено е запазена техническа информация, необходима за по-нататъшната обработка на изображенията и последващи компютърни симулации. На фигура 2а е показано двуизмерно томографско изображение от пациент с рак на млечната жлеза. На същата фигура е показано мястото на лезията (фигура 2б), както и самата лезия (фигура 2в).

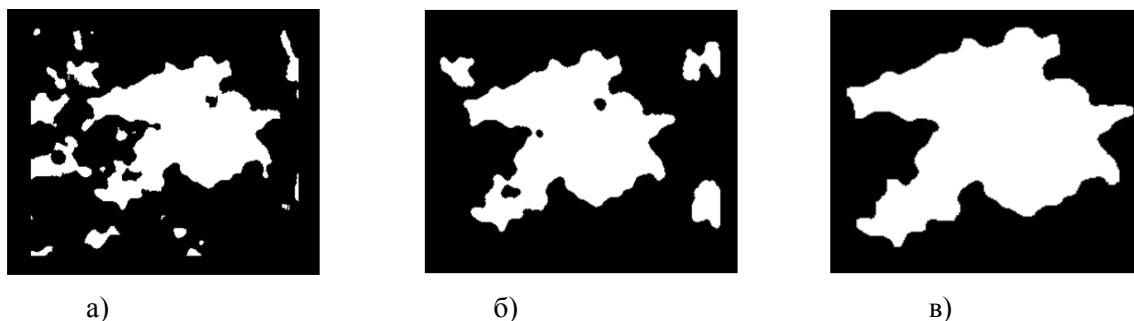


**Фиг. 2.** Пациентски данни: а) двуизмерно томографско изображение от пациент, б) селектирана лезия, в) лезия



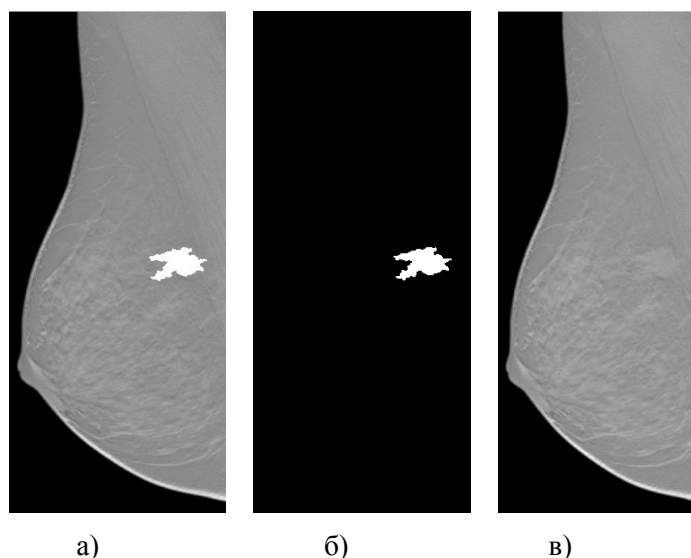
## 2.2. Сегментиране на лезии

Сегментацията на лезиите е извършена със софтуерна платформа, базирана на полуавтоматичен алгоритъм [2]. Създаденият алгоритъм изисква намеса от оператор. Това се състои в избор на изображенията, върху които ще бъде извършена сегментацията, както и областта, където се намира туморното образувание. След автоматичното сегментиране платформата предоставя възможност на оператора за корекции. Самата процедурата по сегментация на туморните образувания се състои от пет основни стъпки, три от които са показани на фигура 3: а) зареждане на клиничните данни, б) предварителна обработка на данните, в) сегментиране на туморните маси, г) потвърждение и/или корекция от оператор, д) създаване на триизмерна вокселна матрица съдържаща сегментирания тумор.



**Фиг. 3.** Етапи при сегментиране на туморни лезии: а) сегментирани елементи, б) премахване на артефакти, в) прилагане на свързващ алгоритъм

На фигура 4 е показан крайният резултат от сегментирането на лезията от пациентски томографски данни.



**Фиг. 4.** Създаване на комбинирано изображение: а) оригинално томографско изображение, б) сегментирана лезия, в) комбинирано изображение

## 2.3. Получаване на различни видове туморни образувания

### 2.3.1. Набор от морфологични операции

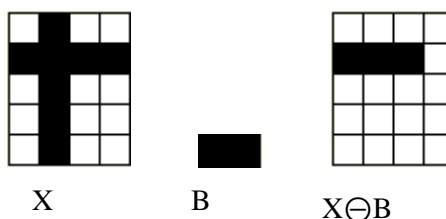
В обработката на изображенията морфологичните операции се използват за отстраняване на пренесен шум от първоначалните изображения в бинарните, т.е. свързване на прекъснати контури, консолидиране на области, изчистване на фонове области и за

откриване на елементи в бинарните изображения на базата на прилагане на логически функции. Тези операции се прилагат върху бинарни изображения с помощта на структурен елемент. Стойностите на пикселите в елемента може да са само единици или единици и нули. Едни от най-разпространените морфологични операции са ерозия и дилатация [3].

А) Ерозия. При ерозията структурният елемент е съставен само от единици, а логическата функция е „И“. Даден пиксел в изходното изображение се установява в единица, ако всички резултати от прилагането на функция „И“ между пикселите на изображението и тези на структурния елемент са единици. Ако с  $X$  означим оригиналното изображение, а с  $B$  структурният елемент, то ерозията може да бъде описана като сечение на всички трансляции на  $X$  спрямо вектора  $b$  (1) :

$$X \ominus B = \bigcup_{b \in B} X_{-b} \quad (1)$$

На фигура 5 е демонстрирана операция ерозия със структурен елемент ( $B$ ) линия.

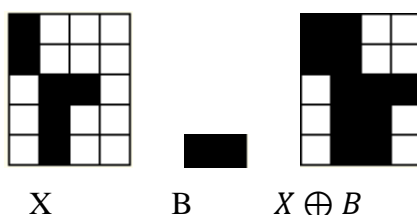


**Фиг. 5.** Морфологична операция ерозия, реализирана между оригиналното изображение  $X$  и структурен елемент  $B$ , с форма на линия

Б) Изглаждане (дилатация). При дилатацията структурният елемент е съставен само от единици, а логическата функция е „ИЛИ“. Даден пиксел в изходното изображение се установява в единица, ако поне един от резултатите от прилагането на функция „ИЛИ“ между пикселите на изображението и тези на структурния елемент са единици. Дилатацията може да бъде описана като обединение на множествата от точки в  $X$  и  $B$  (2):

$$X \oplus B = \bigcup_{b \in B} X_b \quad (2)$$

На фигура 6 е демонстрирана операция дилатация със структурен елемент линия.



**Фиг. 6.** Морфологична операция дилатация, реализирана между оригиналното изображение  $X$  и структурен елемент  $B$ , с форма на линия

С) отваряне, затваряне. Отварянето е резултат от последователно прилагане на ерозия и дилатация с един и същ структурен елемент (3).

$$X^\circ B = (X \ominus B) \oplus B \quad (3)$$

Затварянето е резултат от последователно прилагане на дилатация и ерозия с един и същ структурен елемент (4).

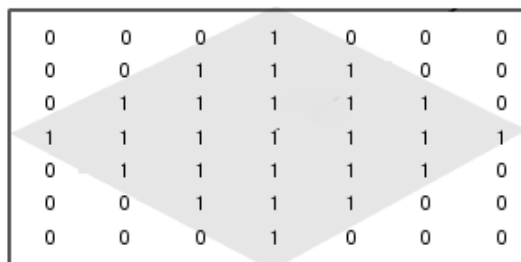
$$X \cdot B = (X \oplus B) \ominus B \quad (4)$$

### 2.3.2. Получаване на туморни образувания

Две от предложените морфологични операции: ерозия и дилатация са приложени върху сегментираните тумори от пациентски данни. За целта са използвани два структурни елемента, показани на фигура 7 и фигура 8.



Фиг. 7. Структурен елемент линия (5x5)



Фиг. 8. Структурен елемент диамант (7)

Структурният елемент линия е с дължина 5 пиксела и с ъгъл от 5 градуса, измерени обратно на часовниковата стрелка, спрямо хоризонталната ос. Структурният елемент диамант е с размер от връх до връх 7 пиксела. Морфологичните операции са приложени последователно срез по срез.

### 2.3.3. Резултати

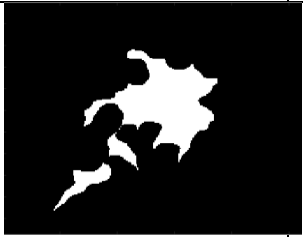
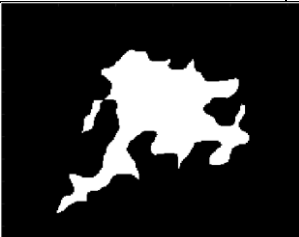

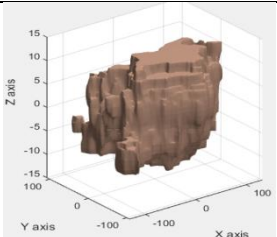

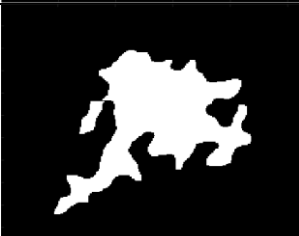
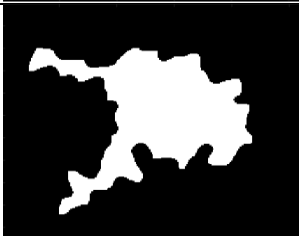
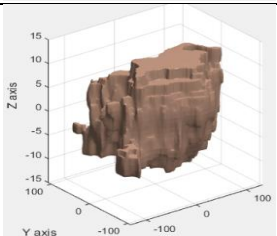



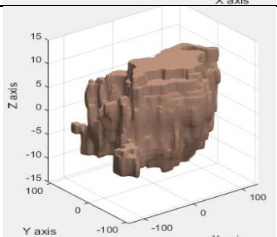
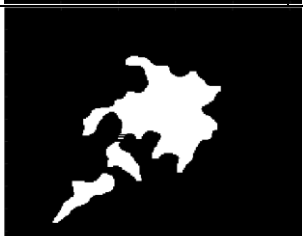
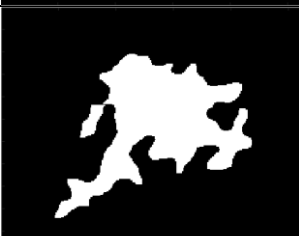
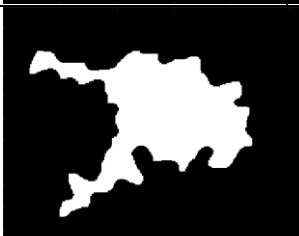
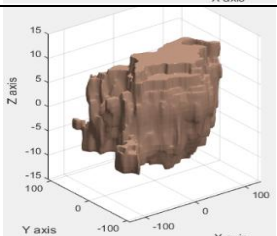
Резултатите от прилагането на морфологичните операции са представени за един от 30те сегментирани тумора. Таблица 1 обобщава следните последователни действия:

1. Отделяне на различните срезове на 3D изображението.
2. Прилагане на морфологична операция ерозия върху оригиналното изображение.
3. Повторно прилагане на морфологична операция ерозия върху оригиналното изображение.
4. Прилагане на морфологична операция дилатация върху повторно ерозираното изображение (изображението от т.3).
5. Повторно прилагане на морфологична операция дилатация върху повторно ерозираното изображение (изображението от т.4).
6. Прилагане на морфологична операция ерозия след повторното прилагане на операция дилатация.




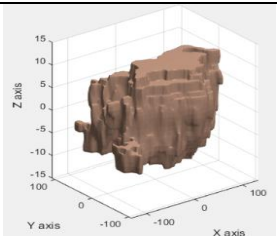



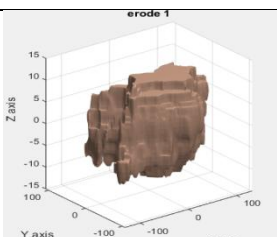
**Таблица 1:** Прилагане на различни морфологични операции върху оригиналното изображение, използвайки структурен елемент линия (5x5) (фигура 7)

	Срез 1	Срез 2	Срез 3	3D изображение
Оригинално изображение				
Ерозия 1				

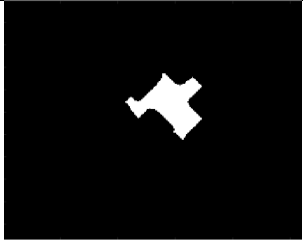


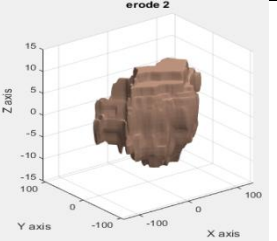



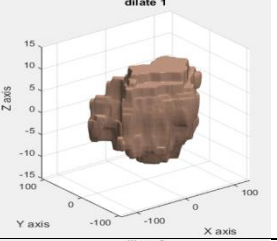

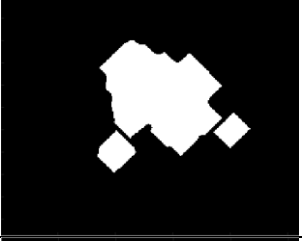
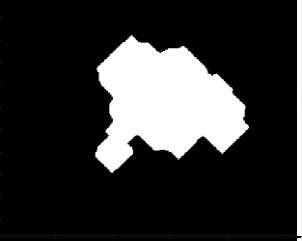
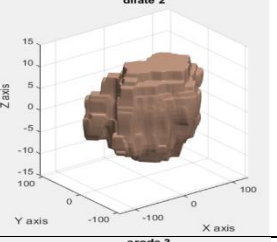



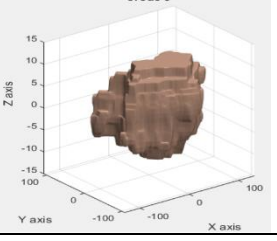
**Таблица 1:** Прилагане на различни морфологични операции върху оригиналното изображение, използвайки структурен елемент линия (5x5) (фигура 7) - продължение

	Срез 1	Срез 2	Срез 3	3D изображение
Ерозия 2				
Дилатация 1				
Дилатация 2				
Ерозия 3				

**Таблица 2:** Прилагане на различни морфологични операции върху оригиналното изображение, използвайки структурен елемент диамант (7) (фигура 8)

	Срез 1	Срез 2	Срез 3	3D изображение
Оригинално изображение				
Ерозия 1				

**Таблица 2:** Прилагане на различни морфологични операции върху оригиналното изображение, използвайки структурен елемент диамант (7) (фигура 8) - продължение

	Срез 1	Срез 2	Срез 3	3D изображение
Ерозия 2				
Дилатация 1				
Дилатация 2				
Ерозия 3				

### 3. Заключение

Прилагането на морфологични операции върху първоначално сегментираните изображения водят до модифициране на последните и създаване на туморни модели от нов вид. Планирани са следните задачи, които за залежали в по-нататъшното развитие на разработката: (а) по-задълбочено изследване на влиянието на формата на структурния елемент върху формата на получената лезия; (б) създаване на триизмерен структурен елемент и прилагане на морфологичните операции с него върху оригиналните триизмерни тумори; (в) анотация на получените туморни образувания.

### Литература

- [1] Национален раков регистър, <https://www.sbaloncology.bg>
- [2] Dukov N., Bliznakov Z., Buliev I., Bliznakova K. (2019) *Creation of computational breastphantoms with extracted abnormalities from real patient images*. In: Lhotska L., Sukupova L., Lacković I., Ibbott G. (eds) World Congress on Medical Physics and Biomedical Engineering 2018. IFMBE Proceedings, vol 68/1. Springer, Singapore.
- [3] Булиев, И., Близнакова, К., Обработка на медицински изображения, ръководство за лабораторни упражнения, ISBN 978-954-20-0773-9, 2017.

- [4] Gonzalez, R. C., R. E. Woods, and S. L. Eddins, Digital Image Processing Using MATLAB, Gatesmark Publishing, 2009.
- [5] Haralick, Robert M., and Linda G. Shapiro, Computer and Robot Vision, Vol. I, Addison-Wesley, 1992, pp. 158-205.

**За контакти:**

ас. инж. Гергана В. Спасова  
катедра „КНТ”  
Технически университет-Варна  
E-mail: g.spasova@tu-varna.bg

ас. инж. Николай Т. Дуков  
катедра „КНТ”  
Технически университет-Варна  
E-mail: nikolay.dukov@tu-varna.bg

доц. Кристина С. Близнакова  
катедра „СИТ”  
Технически университет-Варна  
E-mail: kristina.bliznakova@tu-varna.bg



# ИЗСЛЕДВАНЕ НА АЛГОРИТМИ ЗА ОБХОЖДАНЕ НА ОБЕКТИ, РАЗПОЛОЖЕНИ В ЗОНА ЗА СИГУРНОСТ

Илиян Ж. Бойчев

**Резюме:** В статията е изследвана приложимостта на два алгоритъма за намиране на оптимален път в граф, който се използва за определяне на маршрут на дрон в област на приложение – сигурност. За тази цел се съставя граф от т. нар. waypoints, които представляват GPS координати, които определят маршрута в тримерното пространство. Графът се представя чрез матрица на съседство. Представените алгоритми се разглеждат за три топологии на свързване на обектите.

**Ключови думи:** безпилотен летателен апарат, дрон, граф, gps координати, сигурност

## Research of algorithms for observation the objects in a secure zone

Iliyan Zh. Boychev

**Abstract:** This paper investigates the applicability of two algorithms for routing the optimal path in a graph, used for navigation of a UAV drone in a security setup. For this purpose we make use of a graph of waypoints, which is based on GPS coordinates of flight path maneuver points in the three dimensional space. The graph is represented by a dynamic adjacency matrix, on which we carry out a comparative evaluation of the predicted drone route. The both algorithm are investigated for three topologies for connecting the objects.

**Keywords:** UAV, drone, graph, GPS points, security

## 1. УВОД

Безпилотните летателни апарати (UAVs), намират все повече приложения в нашето ежедневие и тази тенденция ще се ускорова в следващите години. Огромното търсене на UAV-базирани услуги мотивира усилията в световен мащаб за научноизследователска и развойна дейност за разширяване на функционалността и приложимостта на дроновете.

Дроновете се използват в различни области като:

- разпознаване на обекти – разпознаване в изображение, в реално време, следене на подвижни обекти и т.н. [1], [2], [3];
- изследвания в областта на управление на дроновете и създаване на математически модели [4];
- надеждност и сигурност [5], [6];
- предсказване възникването на различни природни явления, мигриране на животни, възникване на пожари др. [7];
- ежедневието (фотография, заснемане на различни събития, транспорт на стоки и др.).

В това изследване дроновете се използват в областта на сигурността.

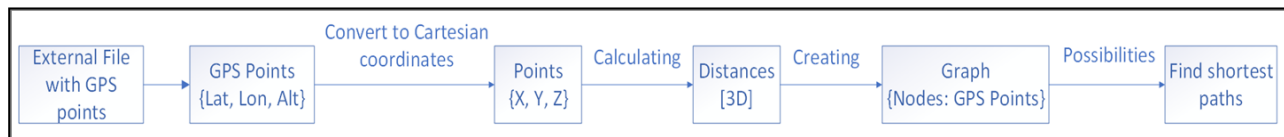
## 2. ИЗЛОЖЕНИЕ

За осигуряване на сигурност и защита е необходимо да се определи маршрут за движение на дрона във въздушното пространство. Този маршрут трябва да бъде оптимален и да отговаря на различни критерии: най-кратък път по разстояние (пряко свързано с консумацията на батерията); път, покриващ максимална зона на видимост на обектите; максимално бързо откриване на евентуален пробив или проблем. В тази статия се изследват два алгоритъма за намиране на оптимален път: алгоритъм на Дийкстра – за намиране на



кратък път по разстояние, и алгоритъм по минимален брой възли. Двата алгоритъма се изследват при използване на три различни топологии на свързване на обектите в зоната за сигурност.

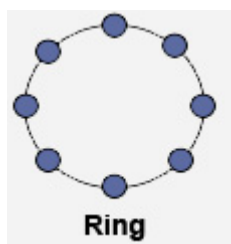
За да бъдат изградени топологиите, е необходимо да се създаде граф, състоящ се от GPS координати (т.н. waypoints), като връзките между точките са разстоянията в 3D пространството (фигура 1). Процесът на преобразуване на координатите, изчислението на разстоянието и създаване на графа са представени в [8].



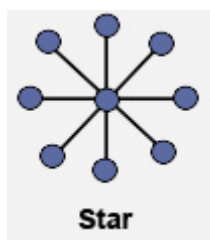
**Фиг. 1.** Създаване на граф в 3D пространството

### Използвани топологии (фигура 2)

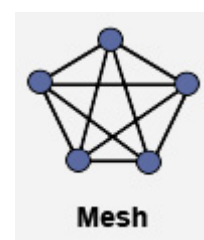
Използват се три топологии на свързване на обектите (сградите).



**а)** тип Кръг



**б)** тип Звезда



**в)** тип Пълна свързаност

**Фиг. 2.** Видове топологии

При топология кръг (Ring) (фигура 2а) обектите са свързани последователно, като всеки има връзка с предишен и следващ. При топология Звезда (Star) (фигура 2б) един от обектите е главен и се свързва с всеки от останалите. При топология пълна свързаност (Mesh) (фигура 2в), всеки обект има връзка с всички останали.

На фигура 3 са представени топологиите, изградени за реалната експериментална обстановка – базата на Технически университет-Варна. Избрани са пет сгради от базата на ТУ-Варна: М (Машинен факултет), ТВ (Топла връзка), НУК (Нов учебен корпус), ЕЛ (Електротехнически факултет) и УПБ (Учебно-производствена база).



**а)** Топология Ring



**б)** Топология Star



**в)** Топология Mesh

**Фиг. 3.** ТУ-Варна топологии



Избраните връзки са на база минимално разстояние между сградите и всяка сграда се свързва с друга с по една линия.

### 3. РЕЗУЛТАТИ

Проведени са няколко експериментални теста с двата алгоритъма за всяка една от топологиите. При всеки тест са избрани различни точки от графа и са съставени пътищата, които се определят от всеки от алгоритмите. За всеки тест са представени графично получените пътища от всеки алгоритъм (фигури 4, 5 и 6), като зелената линия е пътят по алгоритъм на Дийкстра, а бялата линия – по другия алгоритъм. В таблици (Таблица 1, 2 и 3) са описани възлите, техният брой и дължината на всеки път.

3.1. Експериментален тест 1 с начална точка 15(УПБ) – крайна точка 7(ЕЛ) (фигура 4).



а) Топология Ring



в) Топология Mesh



б) Топология Star

Фиг. 4. Експериментален тест 1

Таблица 1. Резултати при избран път 15 - 7

Топология	Дийкстра			Мин. Брой възли		
	Път	Брой възли	Разст.	Път	Брой възли	Разст.
Ring	15 16 3 2 1 0 5 12 11 10 8 7	12	442.429	15 14 20 21 6 7	6	585.834
Star	15 16 3 2 1 0 8 7	8	324.303	15 16 3 2 1 0 8 7	8	324.303
Mesh	15 16 3 2 1 0 8 7	8	324.303	15 14 20 0 8 7	6	448.573

Единствено при топология Star, има съвпадение на пътищата на двата алгоритъма. При топология звезда вторият алгоритъм има два пъти по-малко възела, но общата дължина на пътя е едва около 140m. При втория метод, при топологии кръг и пълна свързаност има еднакъв брой възли, но при втората топология пътят е по-кратък.

3.2. Експериментален тест 2 с начална точка 4(М) – крайна точка 21(НУК) (фиг. 5).



а) Топология Ring



б) Топология Star



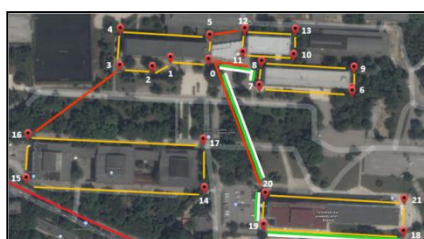
в) Топология Mesh  
Фиг. 5. Експериментален тест 2

Таблица 2. Резултати при избран път 4 - 21.

Топология	Дийкстра			Мин. Брой възли		
	Път	Брой възли	Разст.	Път	Брой възли	Разст.
Ring	4 5 12 11 10 8 7 6 21	9	469.817	4 3 16 15 14 20 21	7	566.241
Star	4 5 0 20 21	5	404.382	4 5 0 20 21	5	404.382
Mesh	4 5 0 20 21	5	404.382	4 5 0 20 21	5	404.382

При вторите две топологии има съвпадение на пътищата и при двата алгоритъма. Пътят при топология Star(Звезда) за алгоритъм на Дийкстра съдържа почти два пъти повече възли в сравнение с останалите два пътя, но дължината е едва около 60м.

### 3.3. Експериментален тест 3 с начална точка 7(ЕЛ) – крайна точка 18(НУК) (фигура 6)



а) Топология Ring



б) Топология Star



в) Топология Mesh  
Фиг. 6. Експериментален тест 3

Таблица 3. Резултати при избран път 7- 18.

Топология	Дийкстра			Мин. Брой възли		
	Път	Брой възли	Разст.	Път	Брой възли	Разст.
Ring	7 6 21 18	4	247.710	7 6 21 18	4	247.710
Star	7 8 0 20 19 18	6	382.628	7 8 0 20 19 18	6	382.628
Mesh	7 6 21 18	4	247.710	7 6 21 18	4	247.710

И за двата алгоритъма и при трите топологии има едни и същи резултати. В този случай няма значение кой алгоритъм ще бъде избран по отношение на резултатния път.

#### 4. ЗАКЛЮЧЕНИЕ

Алгоритъм на Дийкстра - предимства:

- + най-ефективният алгоритъм за най-кратък път;
- + намира всички пътища от даден връх до останалите в графа, което позволява съхраняване и обработка на данните;
- + изключително удобен за използване при дроновете.

Алгоритъм Най-кратък път по минимален брой върхове между два върха - предимства:

- + бърз алгоритъм;
- + подходящ за динамично променящ се граф.

Съвместното използване на двата метода предоставя възможност за създаване на алгоритми за определяне на оптимален път по различни критерии (потенциално опасни зони, труднодостъпни зони, най-често не/използвани маршрути).

Бъдещи насоки за разработка:

- Добавяне на допълнителни обекти за наблюдение;
- Добавяне на нови връзки между отделните обекти;
- Въвеждане на допълнителни критерии за оптималност:
  - време за обхождане;
  - процент покритие на наблюдение при обхождане.
- Въвеждане на различни режими:
  - пълен обход на обектите за максимално покритие на наблюдението;
  - обхождане на потенциално застрашените обекти /области от обекти/;
  - аварийно обхождане при възникнал проблем /неоторизиран достъп в обект/.

#### Литература

- [1] J. Lee, J. Wang, D. Crandall, S. Sabanovic and G. Fox. Real Time Object Detection for Unmanned Aerial Vehicles based on Cloud-based Convolutional Neural Networks, 2017 First IEEE International Conference on Robotic Computing (IRC), 10-12 April 2017
- [2] It Nun Thiang, Dr. Lu Maw, Hla Myo Tun. Vision-Based Object Tracking Algorithm With AR. Drone, International Journal of Scientific & Technology Research Vol. 5, Issue 06, June 2016
- [3] T. Kryjak, M. Komorkiewicz, M. Gorgon. Real-time Implementation of Foreground Object Detection From a Moving Camera Using the ViBE Algorithm. Computer Science and Information Systems 11(4):1617–1637
- [4] Qikai Huang (Bruce Wingo), Mathematical Modeling of Quadcopter Dynamics. Rose-Hulman Institute of Technology
- [5] G. Ngoc; Kwang-Seok Moon; Suk-Hwan Lee; Ki-Ryong Kwon, GIS Map Encryption Algorithm for Drone Security Based on Geographical Features, 2016 International Conference on Computational Science and Computational Intelligence (CSCI), Year: 2016 Pages: 1422 – 1423
- [6] R. Akram; K. Markantonakis; K. Mayes; O. Habachi; D. Sauveron; A. Steyven; S. Chaumette. Security, privacy and safety evaluation of dynamic and static fleets of drones. 2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC) Year: 2017 Pages: 1 - 12
- [7] R. Carlon. Tracking tagged fish using a wave glider. OCEANS 2015 - MTS/IEEE Washington Year: 2015 Pages: 1 – 5
- [8] I. Boychev, Research algorithms to optimize the drone route used for security. XXVII International Scientific Conference Electronics-ET2018, September 13-15, 2018, Sozopol, Bulgaria

**За контакти:**

ас. инж. Илиян Бойчев  
катедра „Компютърни науки и технологии”  
Технически университет-Варна  
E-mail: i.boychev@tu-varna.bg

# МОДЕЛИРАНЕ НА ХАРАКТЕРИСТИЧНА МЕЛОДИЯ НА ПТИЧИ ВОКАЛИЗАЦИИ

Пролет С. Денева, Тодор Д. Ганчев

**Резюме:** Моделирането на мелодията на птичи песни е важна част от анализа и разпознаването на различните видове птици и репертоара им. В настоящата разработка се предлага нов метод за описване на характеристични мелодии в биоакустични записи на птици. За целта първо се прилага метод за предварителна обработка на аудио спектрограмата, който спомага за елиминиране на значителна част от конкурентните акустични събития. Впоследствие, характеристичната мелодия се моделира чрез интерполация на доминантната енергия на сигнала, определена за кратки времеви интервали. Експерименталната валидация на предложения метод за моделиране на характеристичната мелодия е извършена с реални акустични записи на *Myrmotherula multostriata*, регистрирани в техните местообитания в дивата природа.

**Ключови думи:** Цифрова обработка на сигнали, биоакустични сигнали, изчислителна биоакустика, интерполация, характеристична мелодия

## Modeling of characteristic melody of birds vocalizations

Prolet S. Deneva, Todor D. Ganchev

**Abstract:** Bird song melody modeling is an important part of the analysis and recognition of different bird species and their repertoire. In this paper we propose a new method for describing characteristic melodies in bioacoustic bird records. To do this, we firstly apply a method of pre-processing the audio spectrogram, which helps to eliminate a significant part of the competing acoustic events. Subsequently, the characteristic melody is modeled by interpolation of the dominant signal energy determined for short time intervals. The experimental validation of the proposed method of modeling the characteristic melody was performed with real acoustic recordings of *Myrmotherula multostriata* registered in their wildlife habitats.

**Keywords:** Digital signal processing, bioacoustic signals, computational bioacoustics, interpolation, characteristic melody

## 1. Въведение

Регистрирането на биоакустични емисии от ключови животински видове в техните местообитания е силно затруднено поради множеството съпътстващи звукови събития от други видове и шумовете от околната среда. Методите за обработка на акустични сигнали, насочени към потискане на шума във времева или честотна област, не са ефективни, когато съпътстващите звукови събития се припокриват както по време, така и по честота с вокализиациите на изследваните видове. За справяне с този проблем е предложен подход за потискане на смущенията, който работи едновременно върху няколко колони и редове от аудио спектрограмата. Този подход използва методи и алгоритми типични за обработката на изображения. Методите, основани на този подход, се възползват от съпоставянето на времевата и честотната информация за локализация на енергията на сигнала [1]. При тези методи резултатът от обработката на сигнала в определен момент зависи от бъдещите стойности на сигнала, което не позволява реализацията на тези алгоритми в реално време. Записите на биоакустични емисии обикновено се обработват на блокове. Обработка в близко до реално време или в off-line режим е приемлива за голяма част от типичните приложения.

Съществуващите подходи за определяне на характеристиките на птичи вокализации се основават на:

- (1) извличане на описатели на сигнала директно във времевата област;
- (2) получаване на спектрални описатели;
- (3) методи, работещи едновременно във време-честотната област;
- (4) методи използващи други преобразувания на сигнала, с цел извличане на информативни характеристики.

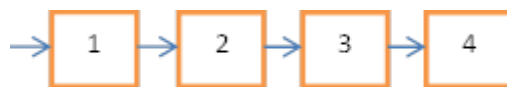
В настоящата разработка е изследвана възможността за определяне на характеристичната мелодия чрез интерполация на основния хармоник на сигнала, определен за кратки времеви интервали. Накратко, в Секция 2 е описано концептуалното решение на проблема и е предложен метод за откриване траекторията на характеристичната мелодия в биоакустични записи на птици, регистрирани в дивата природа. Числени резултати от прилагането на метода са представени в Секция 2, а в Секция 3 е предложено обобщение, изводи и еконкретизирана потенциалната приложимост на метода.

## 2. Метод

Параметричното описание на характеристичната мелодия дава възможност за бързо сравнение между структурите и композициите на множество акустични емисии на пойни птици и други животински видове, които издават звуци с изявена хармонична структура. Основните етапи при моделиране на характеристичните мелодии от аудио записи на птици са:

1. предварителна обработка на сигнала за потискане на шума от околната среда и конкурентни акустични емисии от други животни;
2. сегментиране на сигнала с цел разделяне на обособени акустични събития;
3. интерполация на позициите с доминантна енергия;
4. съставяне на система характеристични уравнения.

Общата структура на концептуалното решение е представено на фигура 1.



**Фиг. 1.** Блокова схема на процеса за откриване на характеристичната мелодия

### 2.1. Предварителна обработка на сигнала

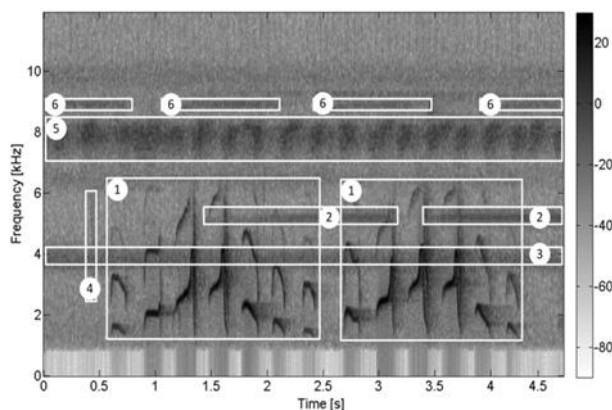
С цел потискане на фоновия шум и определени конкурентни акустични събития, аудио сигналът се подлага на предварителна обработка. В последствие се извършва интерполация за изчисляване траекторията на характеристичната мелодия за обособени акустични събития. В настоящата разработка е използван 2-D метод [1], [4] за предварителна обработка на аудио сигнали, който е подходящ за шумопотискане в записи на птици. На фигура 2 е представен пример за аудио запис на *Myrmotherula multostriata*, направен в естественото местообитание.

На фигура 3 е показан резултатът от прилагането на метода за потискане на фонов шум и нежелани акустични събития [1], приложен върху спектрограмата от фигура 2.

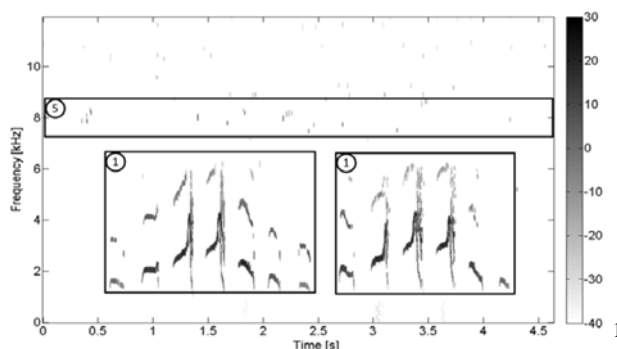
В резултат на обработките, фоновият шум и съпътстващите акустични емисии от други животински видове, различни от Amazonian Streaked Antwren, са подтиснати, с изключение на дребни фрагменти от акустичната емисия, заградена с правоъгълник 5.

За случаите на вокализации с изявени хармонични съставящи има възможност да бъде направена апроксимация на характеристичната мелодия. За целта се прилага маска за отделяне на първия хармоник, последвана от време-честотна локализация на пиковите на енергията. За пример, прилагането на маска и локализацията на пиковите на енергия в изображението от фигура 3 дава възможност за определяне на малък набор от позиции, които

трасират изменението на максималната енергия по време на вокализацията. В последствие локализираните позиции с пикова енергия във време-честотната област се преобразуват до бинарни стойности 0 и 1, което дава възможност за компактно представяне вокализациите и разделянето на сигнала на обособени акустични събития.



**Фиг. 2.** Аудио спектрограма за откъс от пет секунди от запис на *Myrmotherula multostriata* (акустични събития в правоъгълници с етикет 1), но включваща акустични събития от пет други животински вида (заградени в правоъгълници с етикети 2-6)



**Фиг. 3.** Аудио спектрограма след предварителна обработка на сигнала от фигура 2 с 2-D метода [1]

## 2.2 Сегментиране на сигнала

Сигналът, обработен съгласно процедурата, описана в Секция 2.1, се подлага на сегментиране за отделяне на обособени акустични емисии. За целта бинаризираната траектория на пиковата енергия се подлага на конволюция с правоъгълна прозоречна функция с дължина 50 мс. Това позволява елиминиране на малки прекъсвания в траекторията на пиковата енергия, дължащи се на грешки при определянето ѝ, или на наличието на съпътстващи акустични събития от други източници. В резултат на конволюцията участъците, в които липсват изявени вокализации, се преобразуват в дълги последователности от нули. В настоящата разработка е прието акустичните събития, разделени с такива участъци от нули, да се приемат като отделни сегменти, ако нулите са в продължение на повече от 100 мс. Експериментално е установено, че такава граница позволява да се сегментират акустичните вокализации на множество птичи вокализации, в т.ч. и Amazonian Streaked Antwren. В резултат на приложения алгоритъм две акустични

<sup>1</sup> Автор на записа е Jeremy Minns, 2003-04-10, 15:58h, Brazil, Pousada Thaimaçú. Rio São Benedito. Източник: Xeno-Canto.  
URL: <http://www.xeno-canto.org/sounds/uploaded/DGVLLRYDXS/MYRMUL09.mp3>



събития се обособяват като самостоятелни, ако отстоят на повече от 100 мс. Приложен към спектрограмата, показана на фигура 3, така описаният алгоритъм довежда до формиране на 12 обособени сегмента, които се апроксимират самостоятелно, както е описано в следващата секция.

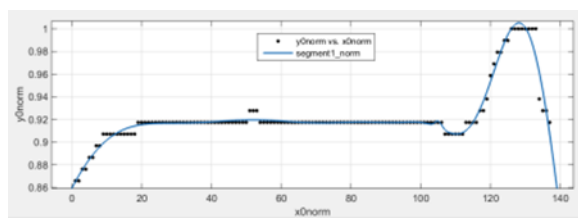
### 2.3. Интерполация

Използвайки позициите с максимална концентрация на енергията, определени съгласно предварителната обработката на сигнала и сегментацията в Секция 2.2, с помощта на числова интерполация се конструират нови числови стойности, които описват обособените участъци от характеристикната мелодия [2].

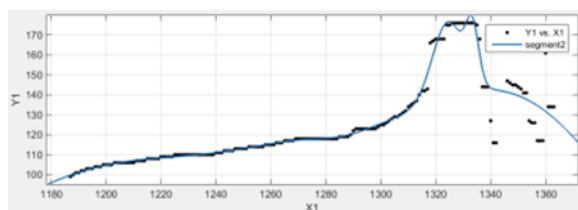
За целта се използва нормалното (Гаусово) разпределение [3]. В общия случай нормалното разпределение се получава от вдигането на квадратна функция в експонента. Това води до класическата камбановидна форма. Коефициентът  $a$  контролира ширината на камбаната,  $b$  контролира централния максимум на камбаната по дължина на оста  $x$ , а  $c$  контролира височината на камбаната. В настоящата разработка се използва модел с шест Гаусови функции:

$$f(x) = a_1 e^{-\left[\frac{(x-b_1)}{c_1}\right]^2} + a_2 e^{-\left[\frac{(x-b_2)}{c_2}\right]^2} + a_3 e^{-\left[\frac{(x-b_3)}{c_3}\right]^2} + a_4 e^{-\left[\frac{(x-b_4)}{c_4}\right]^2} + a_5 e^{-\left[\frac{(x-b_5)}{c_5}\right]^2} + a_6 e^{-\left[\frac{(x-b_6)}{c_6}\right]^2} \quad (1)$$

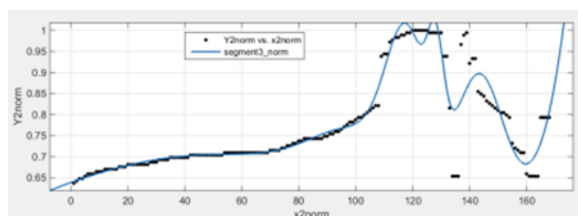
Показаният на фигура 3 сигнал е сегментиран на две големи групи с общо 12 отделни сегмента. За три от тези резултатите от моделирането с (1) са показани съответно на фигура 4, фигура 5, и фигура 6.



Фиг. 4. Апроксимация на характеристикната мелодия за сегмент 1



Фиг. 5. Апроксимация на характеристикната мелодия за сегмент 2



Фиг. 6. Апроксимация на характеристикната мелодия за сегмент 3

Анализирайки времевите сегменти, се получават следните коефициенти на апроксимиращите функции (Таблица 1):

**Таблица 1.** Изчислени стойности на параметрите на модела (1) за три обособени акустични събития от фигура 3

C/S	Segment 1	Segment 2	Segment 3
a1	0,3562(-0,03164;0,7441)	27,06(11,81;42,31)	0,221(0,103;0,3391)
b1	1,561(-1,405;1,718)	1333(1332;1334)	128,1(127,3;129)
c1	0,4019(0,1842;0,6196)	3,275(1,497;5,053)	3,96(2,248;5,671)
a2	0,9161(0,8978;0,9344)	124,5(-193,6;442,5)	0,3478(-0,2405;0,9361)
b2	-0,3891(-1,535;0,7573)	1349(1289;1409)	143,6(140,6;146,6)
c2	3,203(-0,2408;6,647)	61,81(-48,24;171,9)	12,09(2,642;21,55)
a3	0,09361(-0,3748;0,562)	38,18(23,6;52,76)	0,3441(0,06392;0,6243)
b3	-1,66(-2,601;-0,7182)	1324(1322;1326)	118,1(115,9;120,4)
c3	0,65(-0,5028;1,803)	7,897(4,278;11,52)	10,28(4,922;15,64)
a4	0,00593(-0,004227;0,01609)	-2,512(-26,63;21,61)	1,188e+07(-4,849e+10;4,851e+10)
b4	0,9096(0,866;0,9531)	1286(1255;1316)	533,7(-8,574e+04;8,681e+04)
c4	0,03417(-0,03834;0,1067)	13,05(-56,5;82,6)	88,04(-1,027e+04;1,045e+04)
a5	0,07283(-0,2749;0,4205)	20,6(-367,9;409,1)	0,1198(-0,1475;0,3872)
b5	0,7045(-1,433;2,842)	1274(1127;1421)	101,4(63,85;139)
c5	0,5525(-1,029;2,134)	38,76(-160,8;238,3)	22,57(-16,05;61,19)
a6	0,04648(-0,326;0,419)	104,9(50,19;159,6)	0,7045(0,6817;0,7273)
b6	0,9264(0,7548;1,098)	1210(1024;1396)	50,96(8,389;93,53)
c6	0,2903(-0,2067;0,7874)	99,9(-333,4;533,2)	163(20,07;306)

### 3. Заключение

Анализираният метод позволява всеки сегмент от сигнала, съдържащ обособено акустично събитие, да се разглежда като характеристична крива за даден вид вокализация на конкретния биологичен вид. От получените експериментални резултати е очевидно, че при използване на характеристично описание със смес от Гаусови функции се получава добра апроксимация. Относително малката грешка между сигнала и апроксимиращата функция показва, че избраният модел позволява добро приближение до анализирания биоакустичен сигнал. След определяне на коефициентите в модела (1) се съставя системата от характеристични уравнения, описващи група сегменти от сигнала.

Такова компактно описание на акустичните събития създава условия за разработване на нови методи за параметризация на сигнала и за създаване на нови методи за разпознаване на голям брой животински видове по тяхната характеристичната мелодия.

### Литература

- [1] Ganchev T. Computational Bioacoustics: Biodiversity Monitoring and Assessment. First Edition. De Gruyter, Berlin, June, 2017. ISBN: 978-1-61451-729-0, [www.degruyter.com](http://www.degruyter.com)
- [2] Боянов Б. Г., Цифрова обработка на сигнали, Брягпринт АД, Варна, 2003.
- [3] McLachlan, G., D. Peel, Finite Mixture Models, Hoboken, NJ: John Wiley & Sons, Inc., 2000.
- [4] Денева П. Д., Разпознаване на характеристична мелодия на биоакустични сигнали, Акустика 2017, ISSN 1312-4897, стр. 57-60.

#### За контакти:

Доц. д-р инж. Тодор Д. Ганчев  
кафедра „КНТ”  
Технически университет-Варна  
E-mail: [tganchev@hotmail.com](mailto:tganchev@hotmail.com)

инж. Пролет С. Денева  
кафедра “ЕТМ”  
Технически университет-Варна  
E-mail: [pchalakova@mail.bg](mailto:pchalakova@mail.bg)



# AN APPROACH OF BREAST LESIONS MODELLING

Galya Gospodionova

**Abstract:** Development of new detection breast tumor systems relies on the virtual clinical trials, which insist availability of a large number of images with realistic looking pathologies. For this purpose, as realistic as possible models of the breast with lesions are needed. The goal of this study is to create and evaluate a methodology for generation of realistic computational models of breast tumors with irregular shapes and import them into real mammographic images.

**Keywords:** simulation, irregular masses, breast, tumor, mammography, projection images

## 1. Introduction

Breast cancer is the most common heterogeneous malignancy in women under 70 years old [7]. Despite the success in science and medicine the morbidity and mortality still increase. The main factor that improves the chance of survival is the early diagnosis. That is way the new technologies, which aim to detect and diagnose the findings during screening of the breast as earlier as possible, are constantly under development. The collection of a sufficient number of images containing pathological structures is a major limiting factor for performing a clinical study to test a new system [8]. Computerized modelling tools and simulation techniques could create data for these needs and replace expensive conventional clinical trials. The goal of this study is to create and evaluate a methodology for generation of realistic three dimensional (3D) computational models of breast tumors with irregular shapes and import them into real mammographic images. These hybrid images will be stored in an open database, which is intended to be used for the development of new breast cancer detection technologies and performing feasibility studies prior implementation in clinical practice.

## 2. Materials and Methods

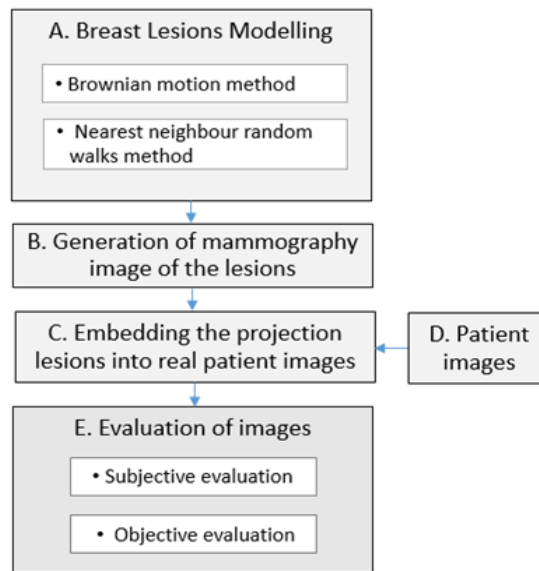
The overall methodology of creation of mammograms with computationally generated breast lesions is schematically shown in figure 1. Each part of the block diagram is presented in the following subsections.

### *A. Methodology for creation of breast lesions*

The methodology for the creation of breast masses with irregular shapes consists of two major steps: (a) use of random walks to create the initial diffusive tumor shape, which is realized by either a Brownian motion or by nearest neighbor random walks, (b) creating of a solid tumor shape by applying a set of 3D filters, as well as morphological operations. In particular, the initial diffusive models were smoothed by applying the following image processing methods: averaging, repeated dilations, morphological opening and closing and final smoothing, all utilized in 3D. Then, the x-ray image of the lesion was generated by a simulation tool. Thereafter, a technique for embedding of the simulated image of the masses in real mammography images was created and applied.

The *Brownian motion random walk* begins by assigning the central pixel of a 3D array - a value of 1 as a tumor pixel [2]. The tumor size is a function of the size of the voxel matrix and the voxel resolution, defined by the user. The user also assigns a number of random walks. Each

random walking process stops either at the matrix boundaries or when the assigned number of steps is reached. The random walk starts from the center of the matrix and each step moves randomly to one of the neighboring voxels, assigning it as the abnormality composition.



**Fig. 1.** An outline of the process of creation of an x-ray mammogram with a breast lesion

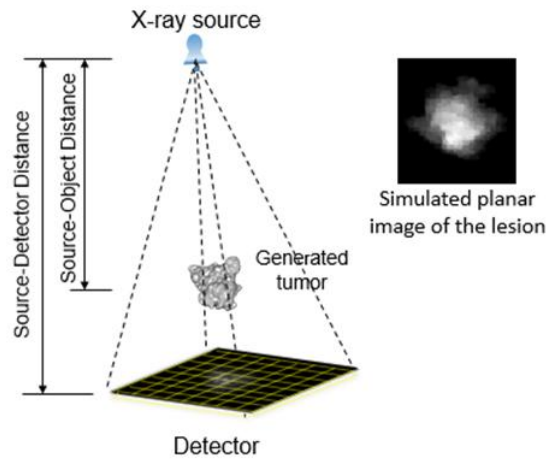
The resulting structure is converted to an abnormality with solid geometry by applying further processing with morphological operations: averaging, dilation, erosion. In these operations, the structuring element is a cube. For instance, the repeating dilation is made with a large cube size (5x5) followed by dilation with a smaller cube size (3x3), while averaging is reached with uniform averaging filter (arithmetic mean). Other morphological image processing methods, closing and smoothing have also cube as a structuring element. Size of the structuring element as well as its shape can be changed.

By changing the number of walks, number of steps per walk, degree of averaging, dilation and erosion, the shape of modelled tumor is also changed as well the malignancy is also altered. Examples of generated abnormalities are shown in figure 3 a, b and c.

*The Nearest neighbor random walk* algorithm is based on the model used by Ruschin et.al. [8]. The random walk begins by assigning a value 1 to the center pixel of the 3D array. For each iteration, the nearest neighbor of the pixel chosen on the last iteration, are randomly selected from a uniform distribution and subsequently a non-zero value is assigned to them as a new tumor pixels. The difference between *The Nearest neighbor random walk* algorithm and *Brownian motion random walk* algorithm is that the first one uses a parameter more – this is the **probability** of the tumor growing in the given direction/walk. On each iteration the amount of the new tumour pixels to be chosen on next iteration is increased with the selected **probability**. The walk is completed when the border of the 3D array is reached [5]. Examples of generated irregular abnormalities are shown on figure 3 d, e and f.

#### *B.Generation of mammography image of the lesions*

X-ray projection images of 3D breast lesions were generated by using an in-house developed software application - *XraySimulator* [1], capable to simulate the x-ray transport through the computational tumors. The geometry and the parameters of the simulated x-ray imaging are shown in figure 2.



**Fig. 2.** Scanning geometry and parameters. Three acquisitions at 0° (CC), 60° (MLO), 90° (ML) were simulated. Source to detector distance is 800 mm, while source to patient table distance is 600 mm. The insert of this figure shows the simulated projection image of the breast lesion

X-ray images were simulated for monochromatic x-ray beams with energy of 20 keV. Distances from the source to the isocenter point, where the centre of the tumour was placed and to the detector surface were 600 mm and 800 mm, respectively. The size of the images was 500 x 500 pixels, while the pixel resolution was 0.1 x 0.1 mm. Analytical relationship between the initial intensity of the x-rays and the intensity registered at the detector is exploited. The transmitted intensity  $I$  reaching the detector pixel is calculated using the Beer's law:

$$I = I_0 * \exp\left(-\int_l \mu(x, y, z) dl\right) \quad (1)$$

where  $\mu(x, y, z)$  is the spatially dependent linear attenuation coefficient,  $l$  is the path length through the object and  $I_0$  is the intensity of radiation at the source that emits to the area of the detector. The generated images are free of scatter, since they exploit the analytical relationships for x-ray matter interaction.

A more complex level of simulation of the radiation interaction at the absorber and the detector includes Monte Carlo techniques. These are used to calculate the photon transport by sampling the interactions with the matter and the distances that x-rays travel until the next interaction. Since these techniques are time consuming, the use of powerful multi-core computers or cloud computing.

Three mammography views were simulated: mediolateral-oblique (MLO), cranio-caudal (CC), and mediolateral (ML) views, which correspond to 60°, 0° and 90°.

### *C. Embedding the projection of the lesions within the patient projection images*

Anonymized planar and free of breast abnormalities patient images obtained with Giotto Tomo IMS system, were used. The pixel size of the images is 100  $\mu\text{m}$  x 100  $\mu\text{m}$ . The created projection image of the breast tumor is then added to the patient image by using the following approach. Initially, patient and lesion images were normalized to their maximal values. Then, the values of the lesion pixels are transferred to the patient images at a position defined by the medical doctors. No further image processing is applied at this stage. The program script implementing the embedding procedure was written in Matlab [9]. Images are then stored in a database and used in the subjective assessment and research and educational activities.

### *D. Evaluation of images*

The realism of the projected breast masses on 2D projection images was evaluated both subjectively and quantitatively. Subjectively, images with embedded projection lesions were

visually assessed by a medical doctor involved in screening and diagnosing of mammography images. The focus of the evaluation is the realism of the lesions: brightness, shape, size and location on patient mammogram. For the objective evaluation, the recently developed software application for quantitative assessment of x-ray images was used [6]. Specifically, the tool is used to compute a set of features from x-ray images such as standard deviation, skewness and kurtosis, fractal and spectrum analysis. These features are then compared to features extracted from real patient images with breast lesions. Pages must not be numbered.

### 3. Results and Discussion

#### A. Created breast lesions

Thirty irregular masses with different sizes and shapes were generated with the proposed two methodology: 15 irregular masses by using the Brownian motion method and 15 such by using the Nearest-neighbour random walks method. The parameters used for their generation are summarized in table 1 (A and B). Each voxel of the three-dimensional array represents an elemental composition, which can be either air (no abnormality) or water. Water was chosen as an elemental composition to represent the mass abnormality, since their x-ray properties are very similar [10].

The data format used for this representation is 16bit unsigned integer. The size of the three-dimensional arrays varied between 100 x 100 x 100 voxels to 200 x 200 x 200 voxels, which corresponded to approximately 15 and 30 MB, respectively. Number of walks and steps varied between 500 to 2000 and 1000 to 5000. These numbers were found to reflect the variety of the shapes which can be obtained with the discussed algorithms. The equations must be centred and numbered - with numbers in brackets aligned to the right.

**Table 1.** Parameters of the tumors generated using Brownian motion method

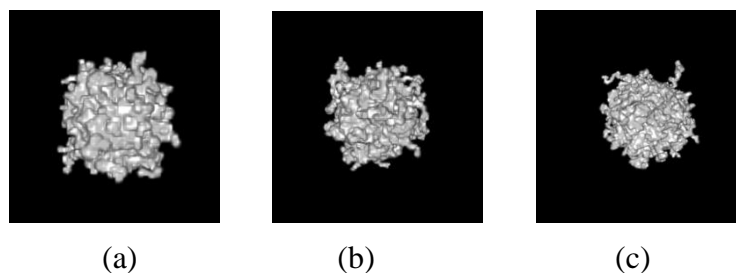
№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Tumor size, pixel	100 x 100 x 100										200 x 200 x 200				
No of walks	500	500	500	500	1000	1000	1000	1500	1500	2000	500	500	1000	1500	2000
No of steps	1000	2000	3000	4000	1000	3000	4000	3000	5000	4000	1000	3000	1000	1000	1000

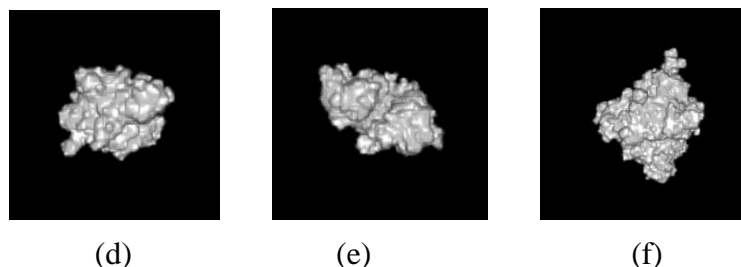
**Table 2.** Parameters of the tumors generated using Nearest Neighbour Random walks method

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Tumor size, pixel	100 x 100 x 100										200 x 200 x 200				
No of walks	500	500	500	1000	1000	1500	1500	2000	2000	2000	2000	500	1000	1500	2000
No of steps	1000	2000	3000	1000	2000	2000	4000	1000	2000	4000	5000	3000	1000	1000	1000

1 pixel = 0.1 mm

Examples of generated models of tumors are shown in Figure 3. As was already mentioned, variation of the parameters in tumor-modeling affects the irregularity of the mass. As shown in the figure 3, examples of masses simulated with nearest neighbor random walk look smoother and more benign, whereas those generated with Brownian motion method look speculated and malignant.





**Fig. 3.** Breast lesion models, generated by (a, b, c) the Brownian motion method; and (d, e, f) by the nearest neighbour random walks method. Specifically, the lesions in (a-c) correspond to lesions 2, 11 and 14 from Table 1, while lesions in (d-f) correspond to lesions 3, 9 and 14 from Table 2

The computational time for the generation of the lesions by the two algorithms was similar. For a tumor array of size 100 x 100 x 100 pixels and 500 walks, the needed computational time was around 5 min., while for a tumor array of 200 x 200 x 200 pixels and 2000 walks this time was approximately 15 min. All simulations ran on a laptop configuration RAM 8GB, Processor Intel(R) Core(TM) i5-32320M CPU @ 2.60 GHz, 64-bit operating system.

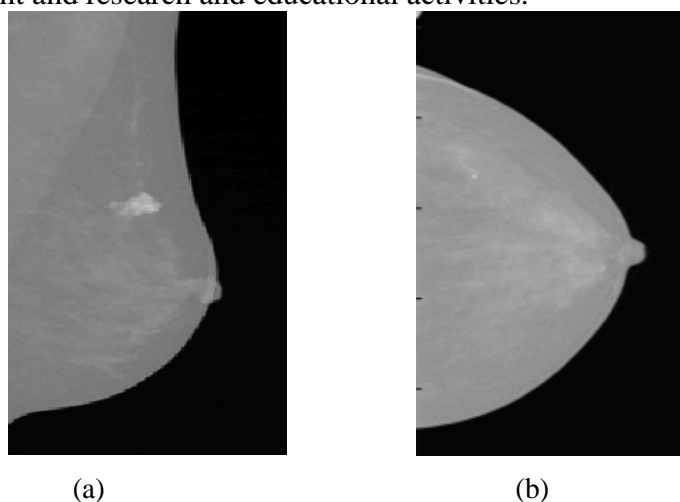
#### A. Mammography images

The created three-dimensional lesions were processed with the *XRAYImagingSimulator*, developed in our Laboratory [2] and mammography images at three different mammographic views ( $60^\circ$ ,  $0^\circ$  and  $90^\circ$ ) were received. Two such examples are shown in Figure 4a, b.



**Fig. 4.** Projection images of two of the generated tumors: (a) Tumor generated with Brownian motion random walk 200x200 pixels, 500 walks and 1000 steps; and (b) Tumor generated with Nearest-neighbour random walks motion random walk 100x100 pixels, 500 walks and 2000 steps.

By using a Matlab script [9], the projection images of the created tumor were mapped to the real mammography images, which are free of lesions. Images are then stored in a database and used in the subjective assessment and research and educational activities.



**Fig. 5.** Real patient images with inserted lesions in (a) MLO view and (b) CC view

Mammograms with the inserted projections of the abnormalities are shown in figure 5a, b. In figure 6a, the projection image of the lesion is inserted into a patient MLO view projection, while in figure 6b the projection image of a small abnormality is inserted into a patient CC view.

#### *A. Evaluation*

The generated thirty mammographic projection of tumour images were visually inspected by a medical doctor with experience in mammography imaging. While the realism of the shapes on the image is quite promising, the comments were mainly to improve the contrast appearance of the abnormality as well as to smooth the tumor outlines. Another limitation of the proposed approaches is the long computational time when high-resolution models are to be generated. This limitation may be overcome with parallel implementation of these algorithms or by using cloud technology.

Objectively, four parameters: skewness, kurtosis, fractal dimension and the power spectrum parameter  $\beta$  were evaluated from these images. The skewness and kurtosis were  $0.14 \pm 0.30$  and  $2.55 \pm 0.29$ , while the evaluated fractal dimension was  $2.60 \pm 0.08$ . The power spectrum parameter  $\beta$  was  $2.79 \pm 0.15$ . These values are well within the ranges of values for these parameters, reported by other researchers [3, 4]. Currently, we are collecting patient images with breast lesions which will be used to evaluate these parameters precisely and perform the comparison with the simulations correctly.

### **4. Conclusions**

The methodology for the creation of breast masses with irregular shapes will be used to generate unique and at the same time realistic in shape and size computational models of breast adenoma, intraductal papilloma, cysts and duct hyperplasia. These computational models are powerful tool in the hands of all professionals working toward the creation of new technology for screening and diagnosing of the breast.

### **5. Acknowledgment**

The methodology for the creation of breast masses with irregular shapes will be used to generate unique and at the same time realistic in shape and size computational models of breast adenoma, intraductal papilloma, cysts and duct hyperplasia. These computational models are powerful tool in the hands of all professionals working toward the creation of new technology for screening and diagnosing of the breast.

### **References**

- [1]. Bliznakova K., I. Sechopoulos, I. Buliev, N. Pallikarakis, "BreastSimulator: A software platform for breast x-ray imaging research", *Journal of Biomedical Graphics and Computing*, 2(1), 2012, pp. 1-14.
- [2]. Bliznakova K., Z. Bliznakov, V. Bravou, Z. Kolotsi and N. Pallikarakis "A 3D breast software phantom for mammography simulation", *Phys. Med. Biol.* 48, 2003, pp. 3699-3721.
- [3]. Byng, J., Boyd, N., Fishell, E., Jong, R., Yaffe, M., Automated analysis of mammographic densities, *Phys. Med. Biol.* 41, 1996, pp. 909-923.
- [4]. Heine, J., Velthuisen, P., Spectral analysis of full field digital mammography data, *Medical Physics*, 29(5), 2002, pp. 3612-4799.
- [5]. Hintsala H., K. Bliznakova, N. Pallikarakis and T. Jämsä, "Modelling of irregular breast lesions," 2009.

- [6]. Marinov St., I. Buliev, L. Cockmartin, H. Bosmans, Z. Bliznakov, G. Mettievier, P. Russo, K. Bliznakova, "Development of a software tool for evaluation of x-ray images: A case study in breast imaging", World Congress on Medical Physics and Biomedical Engineering, IUPESM 2018, June 3-8, Prague.
- [7]. Parkin D. S. , Bray FI, "No Title," *Eur.J Cancer*, vol. 37, 2001, pp. S4-66.
- [8]. Ruschin M, A. Tingberg, M. Båth, "Using simple mathematical functions to simulate pathological structures - Input for digital mammography clinical trial", *Radiation Protection Dosimetry* 114(1-3), 2005, pp. 424-31.
- [9]. MATLAB, <https://www.mathworks.com/products/matlab.html>
- [10]. XCOM: Photon Cross section database: <https://www.nist.gov/pml/xcom-photon-cross-sections-database>.

**For contacts:**

PhD Student, MSc Galya Gospodinova,  
Department of Software and internet technologies  
Technical University of Varna  
e-mail: [genius2604@abv.bg](mailto:genius2604@abv.bg)

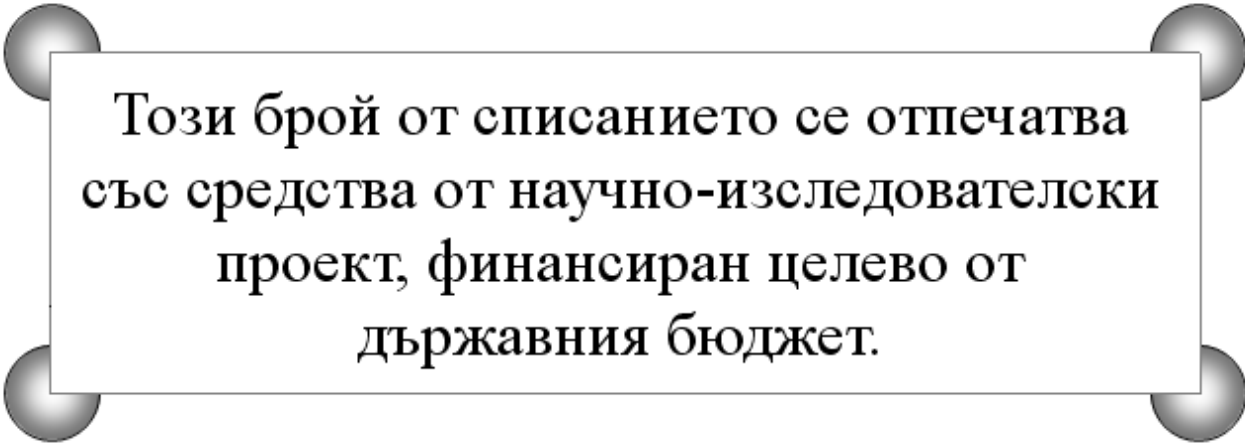


**СОФТУЕРНИ И ИНТЕРНЕТ  
ТЕХНОЛОГИИ**

# ИЗИСКВАНИЯ ЗА ОФОРМЯНЕ НА СТАТИИТЕ ЗА СПИСАНИЕ "КОМПЮТЪРНИ НАУКИ И ТЕХНОЛОГИИ"

- I. Статиите се представят разпечатани в два екземпляра (оригинал и копие) в размер до 6 страници, формат А4 на адрес: Технически университет – Варна, ФИТА, ул. „Студентска” 1, 9010 Варна, както и в електронен вид на имейл адреси: ned.nikolov@tu-varna.bg или yulka.petkova@tu-varna.bg.
  - II. Текстът на статията трябва да включва: УВОД (поставяне на задачата), ИЗЛОЖЕНИЕ (изпълнение на задачата), ЗАКЛЮЧЕНИЕ (получени резултати), БЛАГОДАРНОСТИ към сътрудниците, които не са съавтори на ръкописа (ако има такива), ЛИТЕРАТУРА и информация за контакти, включваща: научно звание и степен, име, организация, поделение (катедра), e-mail адрес.
  - III. Всички математически формули трябва да са написани ясно и четливо (препоръчва се използване на Microsoft Equation).
  - IV. Текстът трябва да бъде въведен във файл във формат WinWord 2000/2003 с шрифт Times New Roman. Форматирането трябва да бъде както следва:
    1. Размер на листа - А4, полета: ляво - 20мм, дясно - 20мм, горно - 15мм, долно - 35мм, Header 12.5мм, Footer 12.5мм (1.25см).
    2. Заглавие на български език - размер на шрифта 16, удебелен, главни букви.
    3. Един празен ред - размер на шрифта 14, нормален.
    4. Имена на авторите - име, инициали на презиме, фамилия, без звания и научни степени - размер на шрифта 14, нормален.
    5. Два празни реда - размер на шрифта 14, нормален.
    6. Резюме и ключови думи на български език, до 8 реда - размер на шрифта 11, нормален.
    7. Заглавие на английски език - размер на шрифта 12, удебелен.
    8. Един празен ред - размер на шрифта 11, нормален.
    9. Имена на авторите на английски език - размер на шрифта 11, нормален.
    10. Един празен ред - размер на шрифта 11, нормален.
    11. Резюме и ключови думи на английски език, до 8 реда - размер на шрифта 11, нормален.
    12. Основните раздели на статията (Увод, Изложение, Заключение, Благодарности, Литература) се формират в едноколонен текст както следва:
      - a. Наименование на раздел или на подраздел - размер на шрифта 12, удебелен, центриран, един празен ред преди наименованието и един празен ред след него - размер на шрифта 12, нормален;
      - b. Текст - размер на шрифта 12, нормален, отстъп на първи ред на параграф – 10 мм; разстояние от параграф до съседните (Before и After) за целия текст – 0.
      - c. Цитиране на литературен източник - номер на източника от списъка в квадратни скоби;
      - d. Текстът на формулите се позиционира в средата на реда. Номерация на формулите - дясно подравнена, в кръгли скоби.
      - e. Фигури - центрирани, разположение спрямо текста: "Layout: In line with text". Номер и наименование на фигурата - размер на шрифта 11, нормален, центриран. Отстояние от съседните параграфи – 6 pt.
      - f. Литература – всеки литературен източник се представя с: номер в квадратни скоби и точка, списък на авторите (първият автор започва с фамилия, останалите – с име), заглавие, издателство, град, година на издаване, страници.
      - g. За контакти: научно звание и степен, име, презиме (инициали), фамилия, организация, поделение (катедра), e-mail адрес, с шрифт 11, дясно подравнено.
- Образец за форматиране можете да изтеглите от адрес <http://cs.tu-varna.bg/> - Списание КНТ, Spisanie\_Obrazec.zip.





Този брой от списанието се отпечатва  
със средства от научно-изследователски  
проект, финансиран целево от  
държавния бюджет.