

Міністерство освіти і науки України  
Харківський національний університет імені В. Н. Каразіна

# ВІСНИК

Харківського національного університету  
імені В.Н. Каразіна

## Серія

«Математичне моделювання.  
Інформаційні технології.  
Автоматизовані системи управління»

**Випуск 47**

Серія заснована 2003 р.

---

# BULLETIN

of V.N. Karazin Kharkiv National University

## Series

«Mathematical Modeling.  
Information Technology.  
Automated Control Systems»

**Issue 47**

First published in 2003

Харків  
2020

Статті містять дослідження у галузі математичного моделювання та обчислювальних методів, інформаційних технологій, захисту інформації. Висвітлюються нові математичні методи дослідження та керування фізичними, технічними та інформаційними процесами, дослідження з програмування та комп'ютерного моделювання в наукоємних технологіях.

Для викладачів, наукових працівників, аспірантів, працюючих у відповідних або суміжних напрямках.

Наказом Міністерства освіти і науки України від 17.03.2020 № 409 наукове фахове періодичне видання Вісник Харківського національного університету імені В.Н. Каразіна серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління» включено до Категорії «Б» Переліку наукових фахових видань України за наступними спеціальностями: 113 – Прикладна математика; 122 – Комп'ютерні науки та інформаційні технології; 123 – Комп'ютерна інженерія; 125 – Кібербезпека.

Затверджено до друку рішенням Вченої ради Харківського національного університету імені В. Н. Каразіна (протокол № 14 від 28.09.2020 р.)

### **Редакційна колегія:**

**Азаренков М.О. (гол. редактор),**

д.ф.-м.н., академік НАН України, проф., ІВТ ХНУ імені В.Н. Каразіна

**Жолткевич Г.М. (заст. гол. редактора),** д.т.н., проф. ФМІ ХНУ імені В.Н. Каразіна

**Лазурик В.Т. (заст. гол. редактора),** д.ф.-м.н., проф., ФКН ІВТ ХНУ імені В.Н. Каразіна

**Споров О.С. (відповідальний секретар),** к.ф.-м.н., доц. ФКН ІВТ ХНУ імені В.Н. Каразіна

**Замула О. А.,** д.т.н., доц., ФКН ІВТ ХНУ імені В.Н. Каразіна

**Золотарьов В.О.,** д.ф.-м.н., проф., ФТІНТ імені Б.І. Веркіна НАН України

**Куклін В.М.,** д.ф.-м.н., проф., ФКН ІВТ ХНУ імені В.Н. Каразіна

**Мацевитий Ю.М.,** д.т.н., академік НАН України, проф., фізико-енергетичний ф-т ХНУ імені В.Н. Каразіна

**Рассомахін С. Г.,** д.т.н., доц., ФКН ІВТ ХНУ імені В.Н. Каразіна

**Руткас А.Г.,** д.ф.-м.н., проф., ФМІ ХНУ імені В. Н. Каразіна

**Стервєдов М.Г.,** к.т.н., доц., ФКН ІВТ ХНУ імені В.Н. Каразіна

**Толстолузька О. Г.** д.т.н., с.н.с., доц., ФКН ІВТ ХНУ імені В.Н. Каразіна

**Угрюмов М. Л.,** д.т.н., проф., ІВТ ХНУ імені В.Н. Каразіна

**Целуйко О.Ф.,** к.ф.-м.н., проф., ІВТ ХНУ імені В.Н. Каразіна

**Шейко Т.І.,** д.т.н., проф., фізико-енергетичний ф-т ХНУ імені В.Н. Каразіна

**Шматков С. І.,** д.т.н., проф., ФКН ІВТ ХНУ імені В.Н. Каразіна

**Щербина В.А.,** д.ф.-м.н., проф., ФМІ ХНУ імені В.Н. Каразіна

**Раскін Л.Г.,** д.т.н., проф., Національний технічний університет "ХПІ"

**Стрельникова О.О.,** д.т.н., проф. Ін-т проблем машинобудування НАН України

**Соколов О.Ю.,** д.т.н., проф., кафедра прикладної інформатики, університет імені Миколая Коперника, м. Торунь (Польща)

Prof. **Harald Richter**, Dr.-Ing., Dr. rer. nat. habil. Professor of Technical Informatics and Computer Systems, Institute of Informatics, Technical University of Clausthal, Germany

Prof. **Philippe Lahire**, Dr. habil., Professor of computer science, Dep. of C. S., University of Nice-Sophia Antipolis, France

**Адреса редакційної колегії:** 61022, м. Харків, майдан Свободи, 6, ХНУ імені В. Н. Каразіна, к. 534.

Тел. +380 (57) 705-42-81, Email: [journal-mia@karazin.ua](mailto:journal-mia@karazin.ua).

**Мова публікації:** українська, англійська, російська.

Статті пройшли внутрішнє та зовнішнє рецензування.

Свідоцтво про державну реєстрацію КВ № 21578-11478 Р від 18.08.2015.

The articles are present research in the field of mathematical modeling and computing methods, information technologies, information security. New mathematical methods of research and management of physical, technical and information processes, research on programming and computer modeling in science-intensive technologies are covered.

For teachers, researchers, graduate students working in relevant or related fields.

By the order of the Ministry of Education and Science of Ukraine from 17.03.2020 № 409 scientific professional periodical Bulletin of V.N. Karazin Kharkiv National University series "Mathematical modeling. Information Technologies. Automated control systems" is included in Category "B" of the List of scientific professional publications of Ukraine in the following specialties: 113 – Applied Mathematics, 122 – Computer Science and Information Technology; 123 – Computer engineering; 125 – Cybersecurity.

Approved for publication by the decision of the Academic Council of V.N. Karazin Kharkiv National University (Minutes № 14 of 28.09.2020).

### **Editorial Board:**

**Azarenkov M.O. (Chief Editor)**, Acad. Of the NAS of Ukraine, Dr. Sc., Prof., HTI V.N. Karazin Kharkiv National University

**Zholtkevich G.M. (Deputy Editor)**, Dr. Sc, Prof. MCS V.N. Karazin Kharkiv National University

**Lazurik V.T. (Deputy Editor)**, Dr. Sc, Prof. CSD HTI V.N. Karazin Kharkiv National University

**Sporov O.E., (Executive Secretary)**, Ph.D. Assoc. Prof, CSD HTI V.N. Karazin Kharkiv National University

**Zamula A.A.**, Ph.D. Assoc. Prof, CSD HTI V.N. Karazin Kharkiv National University

**Zolotarev V.A.**, Dr. Sc, Prof. B. Verkin Institute for Low Temperature Physics and Engineering of the National Academy of Sciences of Ukraine

**Kuklin V.M.**, Dr. Sc, Prof. CSD HTI V.N. Karazin Kharkiv National University

**Matsevity Yu.M.**, Acad. Of the NAS of Ukraine, Dr. Sc., Prof., DPE V.N. Karazin Kharkiv National University

**Rossomakhin S.G.**, Dr. Sc, Prof. CSD HTI V.N. Karazin Kharkiv National University

**Rutkas A.G.**, Dr. Sc, Prof. MCS V.N. Karazin Kharkiv National University

**Styervoyedov N.G.**, Ph.D. Assoc. Prof, CSD HTI V.N. Karazin Kharkiv National University

**Tolstoluzka O.G.**, Dr. Sc, Assoc. Prof. CSD HTI V.N. Karazin Kharkiv National University

**Ugryumov M.L.**, Dr. Sc, Prof. HTI V.N. Karazin Kharkiv National University

**Tseluyko A.F.**, Ph.D. Prof, HTI V.N. Karazin Kharkiv National University

**Sheyko T.I.**, Dr. Sc, Prof. DPE V.N. Karazin Kharkiv National University

**Shmatkov S.I.**, Dr. Sc, Prof. CSD HTI V.N. Karazin Kharkiv National University

**Shcherbina V.A.**, Dr. Sc, Prof. MCS V.N. Karazin Kharkiv National University

**Raskin L.G.**, Dr. Sc, Prof. National Technical University "Kharkiv Polytechnic institute"

**Strelnikova E.A.**, Dr. Sc, Prof., NASU A. Pidgorny Institute of Engineering Problems

**Sokolov O.Yu.**, Dr. Sc, Prof. Nicolaus Copernicus University, Torun, Poland

Prof. **Harald Richter**, Dr.-Ing., Dr. rer. nat. habil. Professor of Technical Informatics and Computer Systems, Institute of Informatics, Technical University of Clausthal, Germany

Prof. **Philippe Lahire**, Dr. habil., Professor of computer science, Dep. of C. S., University of Nice-Sophia Antipolis, France

**Editorial Address:** 61022, Kharkiv, Svobodi sq., 6, V.N. Karazin Kharkiv National University, r. 534.

Phone. +380 (57) 705-42-81, Email: [journal-mia@karazin.ua](mailto:journal-mia@karazin.ua).

**Language of publication:** Ukrainian, English, Russian.

The articles pass internal and external review.

Certificate of state registration: KV № 21578-11478P dated 18.08.2015

## ЗМІСТ

▪ <b>С.І. Богучарський</b> .....	<b>7</b>
Кластеризація колекцій зображень у великих базах даних на основі рекурентної оптимізації	
▪ <b>Р.О. Гамзаєв, М.В. Ткачук, Д.О. Шевкопляс</b> .....	<b>13</b>
Обробка експертних знань в проектуванні лінійок програмних продуктів із застосуванням методу репертуарних сіток	
▪ <b>Н.А. Гук, С.В. Диханов, О.Д. Матющенко</b> .....	<b>25</b>
Алгоритм побудови моделі веб-сайту	
▪ <b>О.І. Кот, І.І. Сватовський</b> .....	<b>35</b>
Аналіз схеми автентифікації на основі використання QR-коду та веб-камери для пристроїв Smart-Mobile	
▪ <b>М.С. Подцикін</b> .....	<b>53</b>
Математична модель профілактики складної технічної системи за станом	
▪ <b>М.О. Слабишев, А.Г. Бердніков</b> .....	<b>72</b>
Модель процесу управління доступом у бездротовій комп'ютерній мережі	
▪ <b>О.Г. Толстолузька, Х.О. Афанасьєва</b> .....	<b>80</b>
Модель урахування дослідницької діяльності працівників системи освіти на базі операційної системи IOS	
▪ <b>О.В. Шарапа, А.Г. Бердніков</b> .....	<b>86</b>
Модель системи управління технологічним процесом в тепличному агропромисловому комплексі	



## CONTENTS

▪ <b>S.I. Bogucharskyi</b> .....	7
Image collections clustering in large databases on the basis of recurrent optimization	
▪ <b>R.O. Gamzayev, M.V. Tkachuk, D.O. Shevkoplias</b> .....	13
Handling of expert knowledge in software product lines development with usage of repertory grids method	
▪ <b>N.A. Huk, S.V. Dykhanov, O.D. Matiushchenko</b> .....	25
Algorithm for building a website model	
▪ <b>O.I. Kot, I.I. Svatovskiy</b> .....	35
Analysis of the authentication scheme based on the use of QR-code and webcam for Smart-Mobile devices	
▪ <b>N.S. Podtcykin</b> .....	53
Mathematical model of condition-based preventive maintenance of a complex technical system	
▪ <b>M. Slabyshev, G. Berdnikov</b> .....	72
Model of control access process in wireless computer network	
▪ <b>O. Tolstoluzka, K. Afanasieva</b> .....	80
Model for taking into account the research activity of education system employees based on the IOS operating system	
▪ <b>A. Sharapa, A. Berdnikov</b> .....	86
Model of process control system in greenhouse agro-industrial complex	

## СОДЕРЖАНИЕ

▪ <b>С.И. Богучарский</b> .....	7
Кластеризация коллекций изображений в больших базах данных на основе рекуррентной оптимизации	
▪ <b>Р.А. Гамзаев, Н.В. Ткачук, Д.А. Шевкопляс</b> .....	13
Обработка экспертных знаний в проектировании линеек программных продуктов с применением метода репертуарных сеток	
▪ <b>Н.А. Гук, С.В. Дыханов, О.Д. Матюшенко</b> .....	25
Алгоритм построения модели веб-сайта	
▪ <b>А.И. Кот, И.И. Сватовский</b> .....	35
Анализ схемы аутентификации на основе использования QR-кода и веб-камеры для устройств Smart-Mobile	
▪ <b>Н.С. Подцыкин</b> .....	53
Математическая модель профилактики сложной технической системы по состоянию	
▪ <b>М.О. Слабишев, А.Г. Бердников</b> .....	72
Модель процесса управления доступом в беспроводной компьютерной сети	
▪ <b>Е.Г. Толстолужская, К.А. Афанасьева</b> .....	80
Модель учета исследовательской деятельности работников системы образования на базе операционной системы IOS	
▪ <b>А.В. Шарапа, А.Г. Бердников</b> .....	86
Модель системы управления технологическим процессом в тепличном комплексе	

УДК 004.932

## Image collections clustering in large databases on the basis of recurrent optimization

S.I. Bogucharskyi

**Bogucharskyi  
Sirhii**

*Candidate of Engineering Science, senior researcher  
V. N. Karazin Kharkiv National University, 4, Svobody Sq., Kharkiv, 61022  
e-mail: [sbogucharskiy@karazin.ua](mailto:sbogucharskiy@karazin.ua)  
<https://orcid.org/0000-0003-4971-4314>*

The following paper considers methods for clustering large amounts of data and proposes a modification of the density-based approach to clustering multimedia objects with disturbance. The analysis of the existing DENCLUE method is carried out, and the matrix influence function is introduced, which makes it possible to effectively use this approach in the analysis of multidimensional objects, the collections of images, video and multimedia data in particular. The introduced matrix form makes it possible to increase the speed of clustering due to the absence of vectorization-devectorization of the initial data.

**Keywords:** clustering, image databases, DENCLUE, influence function.

## Кластеризація колекцій зображень у великих базах даних на основі рекуррентної оптимізації

С.І. Богучарський

**Богучарський  
Сергій Іванович**

*Кандидат технічних наук, старший науковий співробітник  
Харківський національний університет імені В.Н.Каразіна  
площа Свободи, 4, місто Харків, 61022, Україна*

В даній роботі розглянуті методи кластеризації великих об'ємів даних та пропонується модифікація підходу кластеризації мультимедійних об'єктів з збуреннями, заснованого на щільності. Проведено аналіз існуючого метода DENCLUE, та запропонована матрична функція впливу, що дозволяє ефективно використовувати зданий підхід при аналізі багатовимірних об'єктів, в частині, колекцій зображень, відео та мультимедіа даних. Впроваджена матрична форма дозволяє підвищити швидкість кластеризації за рахунок відсутності векторизації-девекторизації вихідних даних.

Дотепер обробка відео викликає цілу низку труднощів, що пов'язані насамперед із розмаїттям тематики, якості та умовами зйомки, для яких неможливо підібрати уніфіковану процедуру розпізнання. Запропонований авторами підхід до обробки відеоданих дозволив виконати скорочення відеороликів та вилучення значущих кадрів, які мають назву ключових кадрів, з урахуванням контенту. Пошук ключових кадрів реалізовано за допомогою математичного апарату діаграм Вороного, які раніше використовувались тільки у сфері геодезії, матеріалознавстві та комп'ютерній графіці для тривимірного моделювання. У статті вирішуються важливі питання щодо пошуку опорних точок (за якими будуються діаграми Вороного) та покращення місця їх розміщення у кольорових зображеннях, якими є відеокадри. Порівняння відеокадрів за допомогою відповідних до них діаграм Вороного надало можливість отримати машинне уявлення про переміщення об'єктів зйомки у просторі та часі.

В статті розглянуто існуючий метод кластеризації мультимедійних даних з підвищеним рівнем шумів. Запропоновано матричний аналог метода кластеризації DENCLUE, призначений для обробки колекцій зображень, збереження у великих базах неструктурованих даних. Алгоритм достатньо простий у чисельній реалізації та характеризується збільшеною швидкістю за рахунок відмови від реалізації допоміжних операцій векторизації-девекторизації вхідних зображень.

**Ключові слова:** кластеризація, бази даних зображень, DENCLUE, функція впливу, матричний аналог, відеокадр.

## Кластеризация коллекций изображений в больших базах данных на основе рекуррентной оптимизации

С. И. Богучарский

**Богучарский  
Сергей Иванович**

*Кандидат технических наук, старший научный сотрудник  
Харьковский национальный университет имени В.Н.Каразина  
площадь Свободы, 4, город Харьков, 61022, Украина*

В данной работе рассмотрены методы кластеризации больших объемов данных и предлагается модификация подхода кластеризации мультимедийных объектов с возмущениями, основанного на плотности. Проведен анализ существующего метода DENCLUE, и введена матричная функция влияния, что позволяет эффективно использовать

данный подход при анализе многомерных объектов, в частности, коллекций изображений, видео и мультимедиа данных. Введенная матричная форма позволяет повысить быстродействие кластеризации за счет отсутствия векторизации-девекторизации исходных данных.

**Ключевые слова:** кластеризация, базы данных изображений, DENCLUE, функция влияния.

## 1. Introduction

The problem of clustering arrays of multidimensional observations is often encountered nowadays and a large number of methods, procedures and algorithms, ranging from purely empirical to strictly mathematical, have been developed for its solution [1-6].

In the most cases, it is assumed that there is a group of  $N$  objects described by  $n$  - dimensional feature vectors  $x(k) \in R^n$ ,  $k = 1, 2, \dots, N$  which must be divided into  $p$  clusters, while this number may be unknown in advance, i.e.  $1 < p < N$ .

Due to the fact that there is no universal algorithm suitable for all possible situations it becomes clear that there is a large number of possible approaches to solving this problem.

A special group of clustering methods is formed by the algorithms designed to process information stored in very-large databases (VLDB) [2, 5], where speed and simplicity of numerical implementation come to the forefront.

In this situation, clustering methods based on the density of data distribution have proven themselves to be quite effective, while the concept of density used here is close in meaning to the distribution density used in probability theory and mathematical statistics. It is the density-based methods that make it possible to form clusters of arbitrary shape when the processed data are distorted by perturbations, and the number of clusters  $p$  is not known in advance. Within the framework of the «density» approach, clusters are understood as areas in the  $n$ -dimensional space of features with a high level of data concentration. These areas are separated by areas with low density and it is here that the disturbances are located.

Thus, algorithms based on the concept of density, in the process of data processing, form areas of arbitrary shape, where the data is concentrated most densely.

The purpose of this work is to analyze density-based clustering methods and develop a modification of the clustering method.

## 2. Existing methods of clustering extremely large amounts of data

The most common method from this class is DBSCAN (Density-Based Spatial Clustering of Applications with Noise), which is computationally simple and resistant to disturbances [7]. The method is based on a number of concepts and definitions, the main of which are internal and boundary points, D-reachability (Density reachability) and D-connectivity (D-connectedness), threshold ( $\varepsilon = Eps$ ) and the minimum number of observations in a cluster (MinPts). In this case, it is assumed that an arbitrary point is directly reachable from any point  $x(q)$  if it is removed in the sense of the accepted metric (traditionally Euclidean) by a distance not exceeding the threshold  $\varepsilon$  which is set a priori. The threshold  $\varepsilon = Eps$  is the main initial parameter of the algorithm set by the user, which is assumed to be a qualified specialist in a specific subject area, in our case, in the field of video processing and computer science.

Based on the selected threshold, the  $\varepsilon$ -neighborhood of the  $x(q)$  point is formed, which consists of all the points that satisfy the inequation

$$\|x - x(q)\| < \varepsilon$$

As for the minimum number of observations in the MinPts cluster, this is a parameter chosen experimentally, usually  $N > \text{MinPts} \geq N+1$ , it is argued that if the  $\varepsilon$ -neighborhood of a  $x(q)$  point contains at least MinPts points, then both  $x(k)$  and  $x(q)$  belong to the same cluster.

If we consider the concept of D-reachability, then a  $x(k)$  point is considered as D-reachable from  $x(q)$  if such «chain» of observations can be formed that each of its elements is directly reachable by its neighbors.

An important factor is that the concept of D-reach is not considered to be symmetric. If  $x(k)$  lies on the cluster boundary, then the symmetry is broken, i.e. this point may contain fewer than MinPts of points in its neighborhood.

It is precisely by finding such boundary points that the formation of clusters is completed. It is clear that in this case it is a priori assumed that the clusters being formed do not intersect. All observations belonging to a specific cluster and having at least MinPts of observations in their neighborhood are called internal points of the cluster. The described asymmetry gives rise to the concept of D-connection, and points  $x(k)$  and  $x(q)$  are called D-connected if they are both reachable from  $x(r)$ , and it is obvious that the concept of D-connection is symmetric.

Based on the introduced concepts, it is possible to define a cluster as a set of D-connected points, and, what is important, this formulation can be extended to other approaches to the clustering problem, where the concept of a metric is used. The clustering process itself can be reduced to a sequence of elementary actions, which, starting from an arbitrary point, finds a set of D-connected data. After all such observations are found, the procedure starts again from an arbitrary previously unanalyzed point and finds all the D-connected data related to it. This happens until all observations of the analyzed group of image objects are exhausted. The set of all objects that are not included in any cluster and contain less than MinPts observations in their neighborhood are treated as noise in the framework of the standard approach, although it may turn out that these points contain unique information that should be carefully analyzed outside of the DBSCAN scope.

It should be noted that the DBSCAN method, due to its simplicity and clarity, has become widespread in many applied problems of data analysis, including segmentation of various kinds of images, where a multidimensional set of features specified in vector form is assigned to each pixel. It is clear that the number of such vectors in the sample can be very large. Of course, some additional characteristics of the analyzed image can be introduced into consideration, however, to successfully solve the problem the user's qualification must be high enough. It is this circumstance, as well as the low level of formalization of this method and the sensitivity to the choice of the algorithm parameters, that gave rise to a number of modifications, devoid of some of the disadvantages of the prototype.

Today, a number of modifications are known, and each new of them sought to minimize the influence of the subjective factor associated with each specific user and additionally formalize the basic procedure.

One of such modifications is DBCLASD (Distribution-Based Clustering of Large Spatial Databases) [8], which can also be used to form clusters of arbitrary shape from "noisy" data. The main advantage of DBCLASD is the ability to process data in a sequential (on-line) mode, while each newly received image can be assigned to one or another cluster based on the analysis of the distributions of distances from the analyzed image to each of the clusters based on the  $\chi^2$ -test. This method has a reduced sensitivity to the choice of the Eps and MinPts parameters, however, it is based on the assumption that the data in each cluster are subject to a uniform distribution law, which is not always the case in real problems, especially those related to image processing.

The development of DBSCAN is also the OPTICS (Ordering Points to Identify the Clustering Structure) algorithm [9], which allows solving clustering problems in conditions when the clusters have not only different shapes, but also different data distribution densities in each class. OPTICS, in addition to the basic concepts and definitions used in DBSCAN, introduces additional characteristics for each observation such as core distance and reachability distance. OPTICS is structurally equivalent to DBSCAN, has advanced functionality, but from a computational point of view, it is much more complex and slower than the prototype, which complicates its use in tasks related to VLDB.

An interesting hybrid of DBSCAN and the popular averages method is Bridge [10], with the help of which the original data array is first processed using the standard averages method, and then DBSCAN is applied to each formed data group, which suppresses noise and restores the data density in each cluster. It is clear that Bridge from a computational point of view is more complex than DBSCAN, however, it is currently used to solve a number of problems related to VLDB [2].

### 3. Clustering based on density

The most formalized and mathematically sound density-based algorithm is DENCLUE (DENSity-based CLUstEring) [11], created for processing large arrays of multimedia data, by forming clusters of arbitrary shape at a high noise level. This method is based on a number of assumptions:

1) the influence of each vector-image on neighboring observations can be formally described by using some function, usually a nuclear one, called the influence function, which describes the relationship of all observations in some neighborhood of the given image;

2) the general density of data distribution in the  $n$ -dimensional space of attributes is formally described as the sum of the influence functions of each observation;

3) clusters are defined as neighborhoods of density attractors (D-attractors), which are, in fact, local maxima of the general data distribution density function.

For some arbitrary point in the feature space, its influence on the image can be described by using the influence function  $f^y(x) = f(x, y)$ , moreover, such functions are most often either a rectangular structure (1)

$$f(x, y) = \begin{cases} 0, & \text{if } D(x, y) > \sigma; \\ 1, & \text{if } - \text{otherwise,} \end{cases} \quad (1)$$

or Gaussian (2)

$$f(x, y) = \exp\left(-\frac{D^2(x, y)}{2\sigma^2}\right) \quad (2)$$

where  $\sigma$  is the parameter of the width of the nuclear function, is the distance, usually Euclidean, between the points  $x$  and  $y$ .

Then, for a set of observations, the general density function can be represented in the form (3)

$$f^x(x) = \sum_{k=1}^N f(x, x(k)) \quad (3)$$

Function (3), is the sum of nuclear functions, characterized by the presence of a set of local extrema-maxima, called D-attractors, each of which represents a separate cluster and can be determined by using one or another optimization procedure. Here we note that the use of the influence function (1) turns DENCLUE into a standard DBSCAN, and if  $f(x, y)$  is continuous and differentiable, such as (2), the standard gradient optimization can be used to find the local maxima. In this case, an arbitrary point is attracted to the D-attractor if the sequence of iterations (4)

$$x^i = x^{i-1} + \eta \frac{\nabla f^x(x^{i-1})}{\|\nabla f^x(x^{i-1})\|}; i=1, 2, \dots; x^0 = -x \quad (5)$$

converges to  $x^*$ .

If  $f^x(x)$  are used as relations (2), (3), then  $\nabla f^x(x) = \sum_{k=1}^N (x(k)-x)f(x, x(k))$ , and the procedure (4) takes the form (5)

$$x^i = x^{i-1} + \eta \frac{\sum_{k=1}^N (x(k)-x^{i-1})f^{x^i}, x(k)}{\left\| \sum_{k=1}^N (x(k)-x^{i-1})f^{x^i}, x(k) \right\|} \quad (5)$$

where  $\eta$  is a search step parameter.

Each of the D-attractors is characterized by its own density function (6)

$$\hat{f}^{x^*}(x) = \sum_{x(k) \in \text{near } x^*} f(x, x(k)) \quad (6)$$

where  $\text{near } x^* = \{x(k): D(x^*, x(k)) \leq \sigma_{\text{near}}\}$  and its extremum determines the coordinates of the cluster centroid.

Of course, from a computational point of view, DENCLA is more complex than any of the algorithms described above, however, its advantages include a high level of formalization, as well as the fact that it generalizes the density-based clustering procedures discussed above.

#### 4. DENCLUE in tasks of clustering image collections

When solving clustering problems, it is always assumed that each multidimensional observation-image is described by a  $n$ -dimensional vector  $x(k)$ , and the entire solution process is associated precisely with vector operations.

In a situation where there is a large collection of images to be clustered, each two-dimensional image must first be vectorized, then the clustering problem is solved, and its result is devectorized, which transforms the vector description into a matrix form. It is possible to significantly simplify the process of clustering arrays without converting them into a vector form, but operating directly with matrices. Thus, the set of initial images is the set of matrices  $x(k)=\{x_{ij}(k)\}$ ,  $x_1 = 1, 2, \dots, m$ ;  $x_2 = 1, 2, \dots, n$ ;  $k = 1, 2, \dots, N$ ,  $x(k) \in R^{m \times n}$ .

Further, instead of the standard vector Euclidean norm its spherical matrix analogue is introduced (7)

$$D_S^2(x,y)=Sp(x-y)(x-y)^T, \tag{7}$$

and the matrix density function

$$f_S^X(x)=\sum_{k=1}^N f_S(x, x(k))$$

In this case, an arbitrary  $(m \times n)$  matrix-image  $x$  is attracted to the matrix D-attractor  $x^*$  if the sequence of iterations of type (4)

$$x^i = x^{i-1} + \eta \frac{\left\{ \frac{\partial f_S^X(x^{i-1})}{\partial x_{ij}} \right\}}{\left( Sp \left\{ \frac{\partial f_S^X(x^{i-1})}{\partial x_{ij}} \right\} \left\{ \frac{\partial f_S^X(x^{i-1})}{\partial x_{ij}} \right\}^T \right)^{\frac{1}{2}}}$$

$i=1, 2, \dots; x^0 = x$

converges to  $x^*$ . Here  $\left\{ \frac{\partial f_S^X(x)}{\partial x_{ij}} \right\}$   $(m \times n)$  matrix, formed by derivatives  $f_S^X(x)$  with respect to the components of the matrix  $x$ .

If the expression (7) is used instead of the matrix function, the optimization algorithm (8) can be rewritten in a simple form

$$x^i = x^{i-1} + \eta \frac{\sum_{k=1}^N r(k,i-1)}{\left( Sp \left( \sum_{k=1}^N r(k,i-1) \right) \left( \sum_{k=1}^N r(k,i-1) \right)^T \right)^{\frac{1}{2}}},$$

where  $r(k,i-1)=(x(k)-x^{i-1})f_S(x^{i-1}, x(k))$ .

Note that this is essentially an extension of (5) to the matrix case.

The use of its matrix analogue instead of a vector description makes it possible to significantly increase the speed of information processing and avoid a number of issues arising in the problem of clustering data described by high-dimensional vectors, which in turn allows processing not only image databases, but also solving problems of clustering video data.

## 5. Conclusion

The article discusses the existing method for clustering multimedia data with a high level of noise. A matrix analogue of the DENCLUE clustering method is introduced, intended for processing image collections stored in large unstructured databases. The numerical implementation of the algorithm is quite simple and its performance is increased due to rejecting auxiliary operations of vectorization-devectorization of the original images.

## REFERENCES

1. Han J., Kamber M. *Data Mining: Concepts and Techniques.*, 2-nd ed., San Francisco: Morgan Kaufmann, 2006., 800 p.
2. Gan G., Ma C., Wu J. *Data Clustering: Theory, Algorithms, and Applications.*, Philadelphia: SIAM, 2007. – 466 p.
3. Abonyi J., Feil B. *Cluster Analysis for Data Mining and System Identification.*, Basel: Birkhäuser, 2007., 303 p.
4. Olson D.L., Dursun D. *Advanced Data Mining Techniques.*, Berlin: Springer, 2008., 180 p.
5. Xu R., Wunsch D.C. *Clustering.*, Hoboken: John Wiley&Sons, 2008., 358 p.
6. Kohonen T. *Self-Organizing Maps.*, 1-st ed., Berlin: Springer, 1995., 501 p.
7. Ester M., Kriegel H.-P., Sander J., Xu X. A density-based algorithm for discovering clusters in large spatial database with noise // *Proc. Int. Conf. on Knowledge Discovery in Databases and Data Mining.*, Portland, Oregon: AAAIO Press, 1996., P. 226-331.
8. Xu X., Ester M., Kriegel H., Sander J. A distribution-based clustering algorithm for mining in large spatial databases // *Proc. 14-th Int. Conf. in Data Clustering “ICDE’98”*, Orlando FLA: IEEE Computer Society, 1998, P. 324-331.
9. Ankerst M., Breunig M., Krilgel H., Sander J. OPTICS: Ordering points to identify the clustering structure // *Proc. 1999 ACM-SIGMOD Int. Conf. Management of Data.* Philadelphia, PA, 1999, P. 49-60.
10. Dash M. “1+1>2”: Merging distance and density based clustering // *Proc. Int. Conf. on Database systems for Advanced Applications.*, Hong Kong. AEEE Computer Society, 2001, P. 30-33.
11. Hu H., Ester M., Sander A. Distribution-based clustering algorithm for mining in large spatial databases // *Proc. 14-th Int. Conf. on Data Clustering “ICDE’98”*, Orlando: FLA AEEE Computer Society, 1998, P. 324-331.

### **Clustering of image collections in large databases based on a recurrent optimization model / S.I. Bogucharskyi // V. N. Karazin Kharkiv National University.**

Approaches to clustering multimedia objects with density-based perturbations are considered. A modification of the existing DENCLUE method has been made, which is based on the introduction of a matrix influence function, which allows using this approach effectively in the analysis of multidimensional objects, in particular, collections of images, video and multimedia. The introduced matrix form allows accelerating the speed of clustering due to absence of vectorization-devectorization of initial data. Refs., 11 titles.



УДК 004.047

## Handling of expert knowledge in software product line development with usage of repertory grid method

R.O. Gamzayev, M.V. Tkachuk, D.O. Shevkoplias

**Gamzayev Rustam**

*PhD, asc. professor, post-doctorate of the Department of Systems and Technology Modeling,  
V.N. Karazin Kharkiv National University, Kharkiv, Freedom Square 4, 61022,  
e-mail: [rustam.gamzayev@karazin.ua](mailto:rustam.gamzayev@karazin.ua) ;  
<https://orcid.org/0000-0002-2713-5664>*

**Tkachuk Mykola**

*Dc. of techn. science, professor, head of the Department of Systems and Technology Modeling,  
V.N. Karazin Kharkiv National University, Kharkiv, Freedom Square 4, 61022,  
e-mail: [mykola.tkachuk@karazin.ua](mailto:mykola.tkachuk@karazin.ua);  
<https://orcid.org/0000-0003-0852-108>*

**Shevkoplias Daria**

*student of the Department of Systems and Technology Modeling,  
V.N. Karazin Kharkiv National University, Kharkiv, Freedom Square 4, 61022,  
e-mail: [shevkoplyasdarya23@gmail.com](mailto:shevkoplyasdarya23@gmail.com);  
<https://orcid.org/0000-0001-5870-771X>*

The development of software product lines (SPL) is a promising way to create competitive software for the modern IT - market. Because of perpetual changes in user needs and the appearance of new software and hardware components for SPL, maintaining a certain level of variability at all stages of their full life cycle (FLC) becomes an actual problem. One of the most important stages is requirements engineering. To support this process it is advisable to utilize the methods of expert knowledge elicitation and analysis. Based on the recognized multilevel approach to SPL engineering proposed by K. Pohl, the research goal is to develop an approach to correct specification of variability features at the initial stages of SPL FLC. The analysis of methods to extract and process the expert knowledge is carried out, their classification is presented, and the method of repertory grid (RG) is chosen for the further usage. It allows analyzing and structuring expert knowledge in a multidimensional information space and makes it possible to consider various contextual factors that affect relevant requirements for the target SPL. The essentials of the RG method are considered and its formal description is elaborated. The examples of its usage in designing “Smart House” systems which allow analyzing the technical, social and economic aspects of user requirements for software and hardware variability at the conceptual design stage of these systems are considered. As the result, a possibility to automate a process of a corresponding FODA-model creation is shown. That provides variability support at the next FLC stages of such systems.

**Keywords:** *expert knowledge, software products line, life cycle, variability, repertory grid method, “Smart-House” system.*

## Обробка експертних знань в проектуванні лінійок програмних продуктів із застосуванням методу репертуарних сіток

Р.О.Гамзаєв, М.В. Ткачук, Д.О. Шевкопляс

**Гамзаєв Рустам  
Олександрович**

*кандидат технічних наук, доцент, докторант кафедри моделювання систем і технологій,  
Харківський національний університет імені В.Н. Каразіна, м. Харків,  
Майдан Свободи 4, 61022*

**Ткачук Микола  
Вячеславович**

*доктор технічних наук, професор, завідувач кафедри моделювання систем і технологій,  
Харківський національний університет імені В.Н. Каразіна, м. Харків,  
Майдан Свободи 4, 61022*

**Шевкопляс Дар'я  
Олександрівна**

*студентка кафедри моделювання систем і технологій,  
Харківський національний університет імені В.Н. Каразіна, м. Харків,  
Майдан Свободи 4, 61022*

Розробка лінійок програмних продуктів (ЛПП) є перспективним шляхом для створення конкурентноздатного програмного забезпечення на сучасному ІТ – ринку. За умов постійних змін потреб користувачів та появи нових програмно-технічних засобів для реалізації таких систем актуальною проблемою стає можливість забезпечення певного рівня варіабельності проектних рішень на всіх етапах їх повного життєвого циклу (ПЖЦ). Одним з найбільш складних з них є етап визначення початкових вимог користувачів до майбутньої системи і для підвищення ефективності цього процесу доцільно застосовувати методи виявлення та аналізу експертних знань у відповідній предметній області, для якої створюється та чи інша ЛПП. В статті, на основі аналізу загально визнаного багаторівневого підходу до розробки ЛПП, запропонованого в роботах К. Поля (К. Pohl), сформульована основна мета дослідження: розробити підхід до коректного визначення властивостей варіабельності цільової системи вже на початкових етапах її ПЖЦ. Для цього проведено аналіз та запропоновано класифікацію методів видобутку та обробки експертних знань, і для подальшого застосування мотивовано обрано метод побудови репертуарних сіток (РС). Він дозволяє проводити аналіз та структурування експертних знань у багатовимірному інформаційному просторі, що дає можливість враховувати різні контекстні чинники, які можуть впливати на множину відповідних вимог до майбутньої системи. Розглянуті концептуальні засади методу побудови РС та запропоноване формальне теоретико-множинне визначення для його основних складових: елементів, конструктів та рейтингових оцінок. На прикладах з предметної області розробки систем «Розумний дім» з використанням такого підходу наведені конкретні приклади побудови РС, які дозволяють враховувати технічний, соціальний та економічний аспекти вимог користувачів щодо підтримки варіабельності програмно-апаратного забезпечення на етапі концептуального проектування цих систем. В результаті дослідження показана можливість реалізації автоматизованого процесу побудови відповідної FODA-моделі для забезпечення властивостей варіабельності на наступних етапах ПЖЦ систем цього типу.

**Ключові слова:** експертні знання, лінійка програмних продуктів, життєвий цикл, варіабельність, метод репертуарних сіток, система «Розумний дім».

## Обработка экспертных знаний в проектировании линеек программных продуктов с применением метода репертуарных сеток

Р.О. Гамзаев, Н.В. Ткачук, Д.А. Шевкопляс

**Гамзаев Рустам Александрович**

*кандидат технических наук, доцент, докторант кафедры моделирования систем и технологий,  
Харьковский национальный университет имени В.Н. Каразина, г. Харьков,  
Площадь Свободы 4, 61022*

**Ткачук Николай Вячеславович**

*доктор технических наук, профессор, заведующий кафедрой моделирования систем и технологий,  
Харьковский национальный университет имени В.Н. Каразина, г. Харьков,  
Площадь Свободы 4, 61022*

**Шевкопляс Дарья Александровна**

*студентка кафедры моделирования систем и технологий,  
Харьковский национальный университет имени В.Н. Каразина, г. Харьков,  
Площадь Свободы 4, 61022*

Разработка линеек программных продуктов (ЛПП) - перспективный путь создания конкурентоспособного программного обеспечения для современного IT-рынка. Постоянные изменения в потребностях пользователей и появление новых средств разработки ЛПП требуют обеспечения свойств вариабельности на всех этапах их полного жизненного цикла (ПЖЦ). Важным этапом ПЖЦ является анализ начальных требований и для поддержки этого процесса рекомендуется использовать методы обработки экспертных знаний. На основе анализа многоуровневого подхода К. Поля к созданию ЛПП определена цель исследования: разработать подход к корректной спецификации вариабельности на начальных этапах ПЖЦ ЛПП. Проанализированы методы извлечения и обработки экспертных знаний, предложена их классификация и для дальнейшего использования выбран метод репертуарных сеток (РС). Он позволяет анализировать и структурировать экспертные знания в многомерном информационном пространстве и учитывает различные факторы, влияющие на требования к целевой ЛПП. Рассмотрены основы метода РС, дано его формальное определение и показаны примеры его использования для проектирования систем «Умный дом», что позволяет учитывать технические, социальные и экономические аспекты требований пользователей к изменчивости программного и аппаратного обеспечения на этапе концептуального проектирования этих систем. В результате показана возможность автоматизации процесса создания соответствующей FODA-модели, обеспечивающей поддержку вариативности на следующих этапах ПЖЦ таких систем.

**Ключевые слова:** экспертные знания, линейка программных продуктов, жизненный цикл, вариабельность, метод репертуарных сеток, система «Умный дом».

### 1. Introduction: Problem actuality and research goals

The current situation in Software Engineering (SE) is characterized by the growing number of sophisticated methodologies, advanced methods and powerful tools aiming to develop software products at a lower cost, in a shorter time, and accounting for their quality attributes. The development of software product lines (SPL) is a highly promising way to create competitive software for the modern IT - market. Because of perpetual changes in user needs and the appearance of new software and hardware components for SPL design and implementation, maintaining an appropriate level of variability at all stages of their full life cycle (FLC) becomes an actual problem. The recognized multilevel approach to SPL engineering within its FLC was proposed by K. Pohl [1], and its graphical representation is shown in Fig.1.

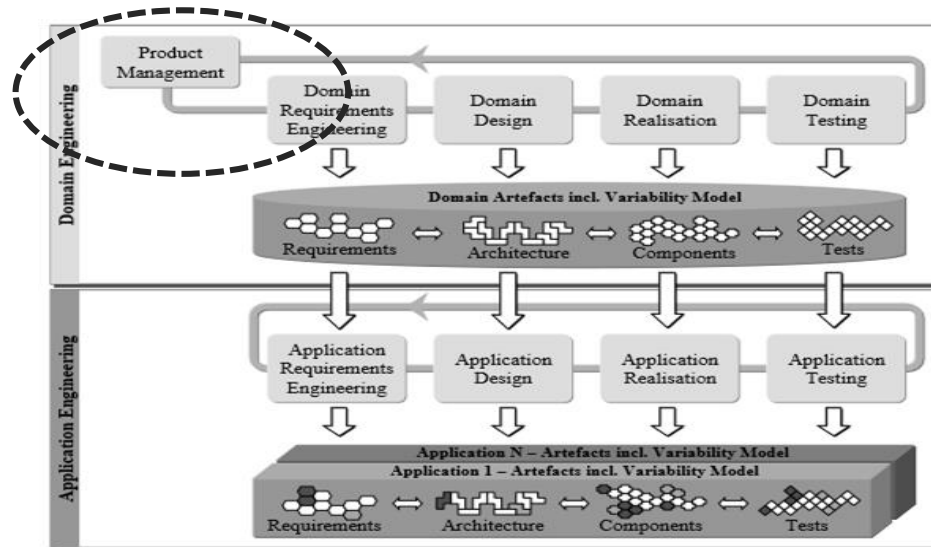


Fig. 1 The software products line engineering framework [3]

According to this engineering framework the FLC of any SPL includes a set of activities to produce the appropriate project artifacts which are structured at the 2 logical levels, namely (see Fig.1):

(I) Domain Engineering Level (DEL) provides a reusable operational platform and, therefore, it defines the *commonality* and the *variability* of future SPL components;

(II) Application Engineering Level (AEL) is responsible for the development of SPL-applications from the operational platform established at the DEL.

Each of these 2 levels: the DEL and the AEL, has a similar sequence of the several ordered project phases: *Requirements Engineering*, *Design*, *Realization* and *Testing*, and in the theory and practice of the modern SE there are a lot of appropriate development methods and CASE-tools to support them effectively [2, 3].

But at the DEL level there is a special project phase (see Fig.1): a phase of *Product Management* which is closely interconnected with a phase of *Domain Requirements Engineering*. Its main goal is eliciting and structuring all important user needs which have to be fulfilled by a target software product, especially, in a SPL to be developed. It should be mentioned that this phase is the most uncertain and weak-formalized one in FLC, and in order to support business analysts and software architects, who are supposed to make the project decisions, some interdisciplinary approaches are to be used (see e.g. in [4]). One of such approaches is a knowledge-based one [5], where a term *knowledge* can be formulated accordingly to the recommendations of the System Thinking World Community [6], namely: a *knowledge* is a collection of structured information objects and relationships, combined with appropriate semantic rules for their handling in order to get new proven facts about a given problem domain. That is why the main goal of the current research is to elaborate an approach to effectively handling the knowledge on the *Product Management / Domain Requirements Engineering* phases within SPL full life cycle (see Fig.1).

This paper is structured as follows: in the second section the variability issues for the one of important modern technical domains: development of smart-house systems (SHS) are discussed; in the

third section some approaches to expert knowledge handling are overviewed and classified, and the method of repertory grid (RG) is chosen for the further consideration; in the fourth section the essentials of the repertory grid method are presented and illustrated with the examples in the SHS domain, and in the fifth section the usage of RG for building of a target feature-oriented model is shown. In the last section the conclusions and application for the further works are discussed.

## 2. The application area and problem statement: variability issues in “Smart-house” system (SHS) domain

In order to present our approach, we have decided to choose SHS which is rather complicated problem area, because the solutions that are developed within this domain usually combine software and hardware parts such as sensors, actuators, smart lighting systems, devices responsible for opening of doors/windows, etc. (Fig. 2). Hardware parts are developed by different vendors and use different protocols and, therefore, could require a lot of proprietary software applications. The system development under these conditions has the following potential drawbacks:

1. SHS management complexity is increased;
2. It becomes very complicated to implement some scenarios that interconnect devices from different vendors;
3. A number of different applications is required in order to work with different parts of the system.

It should be noted that most of the devices have quite similar functionality. For example, an alarm system could send a notification to the messenger, and a water sensor could do the same. In the same way some devices could share information about current conditions. That requires providing similar settings for different devices.

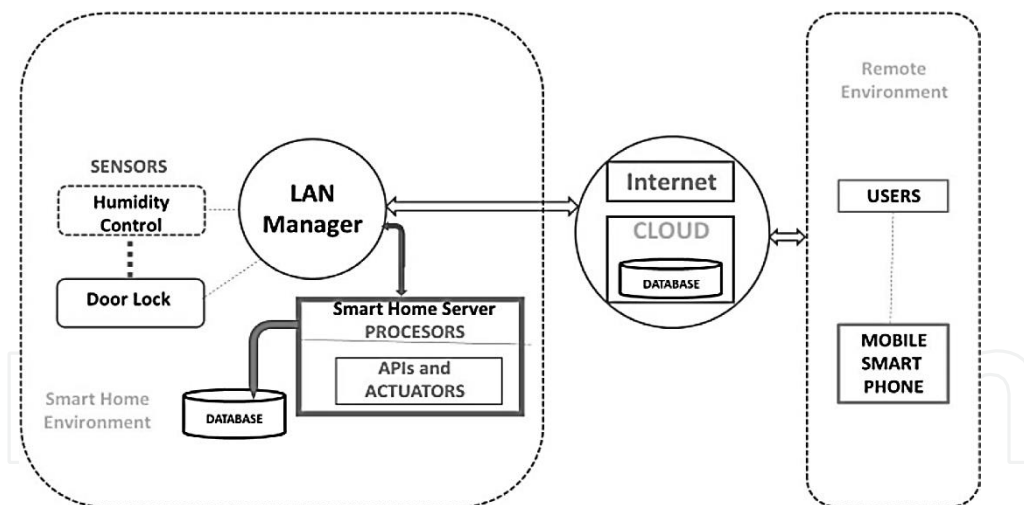


Fig.2 Modern SHS multi-level architecture [7]

There are the following SHS-specific features and problems:

- a lot of interconnected hardware and software components;
- all components have a lot of specific features and parameters;
- components operate in real-time mode;
- operation environment is dynamic and changeable;
- SHS has several groups of users;
- hardware components developed by different vendors;
- different protocols for communication.

After analyzing some of the abovementioned problems it becomes possible to merge them into the single integration problem at the software level. The main issue here is necessity to combine specific components of different platforms into one control system, with the possibility of centralized management of all kinds of equipment. The second issue can be expressed as the need to implement support for variability (or adaptability) of properties of software components, basing on the permanent presence of a set of different groups of requirements and their connection/dependence in the user group.

Based on these conditions and in order to formalize analysis process and synthesis of the SHS system structure, the corresponding initial domain model was built using the FODA notation [8], which is shown in Fig. 3, where the problem part of the system (hardware modules) is outlined with a bold line. The structure of this model contains corresponding sensors and actuators, and different subsystems. Sensors and actuators are controlled by special control unit - "Hub", and system engineers responsible for SHS maintenance, usually need a separate hub for each of the subsystems.

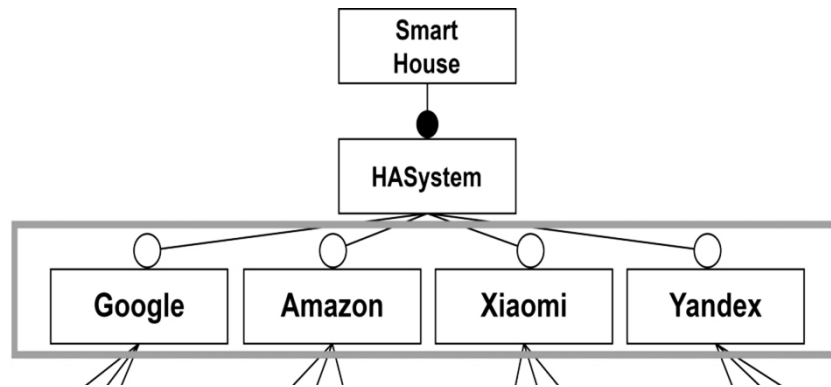


Fig.3 Technological variability issues in SHS

According to this first vision of the FODA-model (Fig.3), the integration problem is the necessity to use separate hub for each vendor, and, often, to use the proprietary applications to control the SHS elements due to the lack of integration and implementation of these functions into some common control system [8]. The second group is partially related to the first one, since a large amount of duplication of different functionality appears at different levels of the system, from the user interface to the devices themselves. When introducing variability support processes, the interface can be combined into a single application, due to the integration of the functionality of the selected components into one unit.

According to the problem mentioned above the main development objective is “How to design and implement configurable and adaptable complex software and hardware solutions, taking into account different users’ needs and system requirements?”

One of the approaches that makes it possible to omit the problem described above and to reach development objective is the SPL development [1, 3]. Any SPL includes methods for the variability management, the main idea is to manage configuration within a product family, plan the product development and elaborate a strategy for the further development.

The SHS SPL development process starts from the requirements elicitation and modeling the features for SPL. Achieved results are used to determine the similarity and variability in the SPL and define dependency in the system [8]. On the next stage of the SHS SPL modeling, previously defined variable and non-variable components become part of the feature model that defines a functional structure, similarity and reuse of the system elements. Dependencies and interfaces of such components are defined on this stage as well.

After domain analysis is finished and conceptual feature model is elaborated, it is used for the product development. The system architecture is build using variability components according to the model.

### 3. Some approaches to handling of expert knowledge and the motivated choice of repertory grid method

The concept of knowledge-based systems is focused on the processing of human knowledge that was received as a result of the experts’ actions in the professional domain [9]. An example of expertise in the SPL domain is the knowledge connected with the software architecture of such systems, conditions and metrics that should be considered, as well as, possible design/architecture patterns.

The difficulties in the knowledge elicitation process result from the following factors:

1. The influence of the experts’ experience or their personal opinion in the selected domain;

2. Absence of the theoretical methods to describe expertise;
3. Difficulty in the verbal expression of the knowledge.

Therefore, the main task could be formulated in the following way: to develop a control system that would facilitate the automated process of expertise elicitation.

Systems that could be used for knowledge elicitation could be classified by the expert methods used as follows [9]:

1. Systems based on Decision Trees;
2. Systems based on the psychological methods, such as Repertory Grid, Cluster Analysis, Multidimensional Scaling;
3. Systems that use models and methods for solving specific problem - Problem-Solving Methods;
4. Systems based on Case-Based Reasoning;
5. Inductive methods of knowledge elicitation, in which knowledge of expertise domain can be represented by examples;
6. Systems with combined methods and approaches for knowledge elicitation.

This classification is represented in Fig. 4:

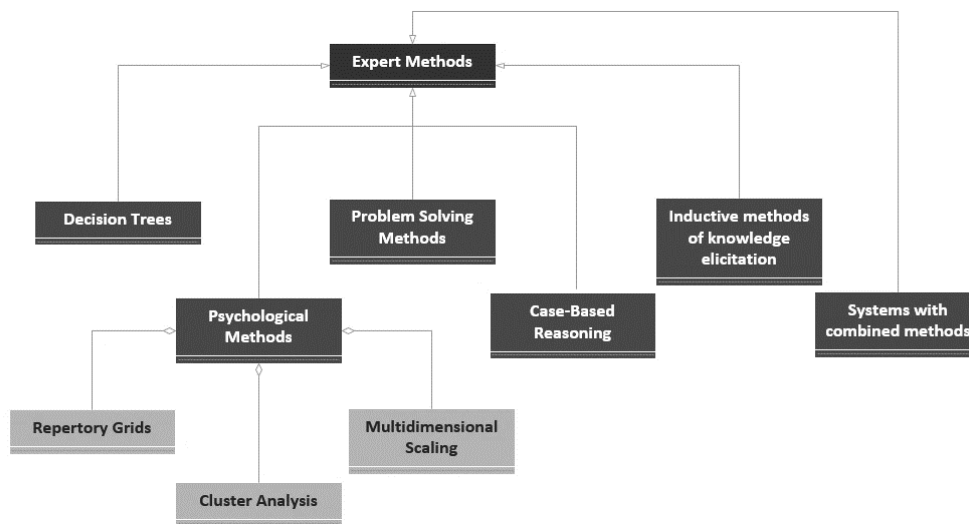


Fig.4 The classification of the expert methods based on their review given in [9]

We will overview some of these methods more closely.

Decision Trees method is the most commonly used and often applied to solve the problem of classification in machine learning. It is relatively simple method of data formalization that represents data as a directed acyclic graph consisting of nodes with rules and edges.

Repertory Grid method is an interview method based on Personal Construct Psychology theory initially designed by George Kelly. In some cases, this method is used together with a cluster analysis for the most effective handling of the expert knowledge. It is used to study personal and interpersonal systems of meaning, conceptual human structure (cognitive inclinations) in terms of interpersonal relationships by classifying a set of certain elements (e.g., other people) in terms of personal constructions [10]. The repertoire grid consists of the following components: elements, concepts and mechanisms that connect these elements and structures. The main idea of this method of identifying expert knowledge is to compare some elements according to the given characteristics.

Problem-Solving Methods are used to extract expert knowledge, build the appropriate database, identify missing knowledge and add/correct the database basing on the method of heuristic classification [9]. The problems solving method (heuristic classification) is performed in the following way:

1. Allocates for each failure or problem an explanation or possible alternatives;
2. Defines information for differentiation of these alternatives;

3. Select an approach to resolve the problem and plan the implementation of the best alternatives.

In this way the knowledge elicited from experts comes as a response to failures and complaints to be resolved. Thus, a multilevel network of problems and possible explanations is formed [9].

Case-Based Reasoning methods are based on the previous experts' problem solving practice and consist of the following stages [9]:

1. Eliciting one or more cases from expert practice;
2. Reusing and revising past knowledge and applying it to the new model, as well as, adapting it to the new situation in some cases;
3. Elicited knowledge is organized with a set of tables in order to consider all possible alternatives of reasoning and present the information as set of rules.

Inductive methods use a set of examples to represent knowledge about the expert domain. Such methods automatically build hypotheses about the connections and patterns of the studied domain. Inductive methods are classified according to the following criteria [9]:

- inductive rules learning,
- data-driven learning,
- model-driven learning,
- heuristic-based learning.

Systems that use a combination of different methods and approaches are the most powerful for expert knowledge elicitation, because shortcomings of one method are compensated by the advantages of others. However, such systems are often difficult to manage and maintain. There are following stages in such methods: knowledge extraction, task decomposition, information processing, step-by-step testing, data types integration, etc. [9].

For the further research the Repertory Grid Method is chosen, and its advantages for this purpose are discussed below.

#### **4. Essentials of the repertory grid method: conceptual approach, formal definition and knowledge structured analysis with the examples in SHS domain**

The main idea of the method of repertory grid (RG) is based on the assumption that a cognitive thinking of experts in some problem domain can help to decrease the process complexity and developer efforts by requirements engineering, especially, to determine appropriate variability features for the software products to be developed [10,11].

Any RG is built as a matrix of some dimension (usually, not higher than 3x3) with different contexts (represented as matrix columns) that influence system functionality in any way, and its characteristics (structured as matrix rows), that can vary depending on a specific problem domain. According to this vision an appropriate RG includes 3 sets of basic components (below they are shown on the example of SHS domain):

- a set of elements: they can be defined based on any domain expert interview which has to be divided into some semantic objects to be presented as RG columns, e.g.: “*Smart TV*”, “*Smart light sensors*” ...);
- a set of constructs: they have to reflect the alternative variable features of RG elements, and they are placed in the RG rows depending on the appropriate context, e.g. “*Smart TV should be turned on / turned off*”); “*Windows should be opened/ be closed*”, ...;
- a set of possible rating values: to evaluate the behavior of the system under the influence of certain elements the evaluation scale is defined (usually in the range from 1 to 5) where the lowest number of range represents the left bipolar construct, and the highest number of range is the right bipolar construct, e.g. the element “*Smart TV*” *should be turned on* (the value of 1) *or turned off* (the value of 5), depending on a given context.

Now we consider the usage of RG method for the handling of expert knowledge in the SHS development domain (see in section 2), basing on the approach elaborated in [N]. For this purpose, we propose a set-theoretical formalization of the RG – framework, therefore any RG can be represented as a tuple:

$$RG = \{E, C, R\}, \quad (4.1)$$

where  $E$ ,  $C$ , and  $R$  are the set of elements, set of constructs and set of rating values, respectively.



As already mentioned, any RG is a special kind of matrix, and due to its multi-dimensional structure it is possible to represent some different aspects of an initial problem domain. Another advantage of RG method is a good visualization of expert knowledge handling. In the case of SHS development (as a typical social-technical system) these aspects (or dimensions) can be defined as [11]:

- social aspects (SHS has to be a user-friendly solution),
- economical aspect (project expenses should be reasonable for SHS owners),
- environmental aspect (any SHS has to be eco-friendly).

Taking into account these aspects the appropriate problem space (PS) for the conceptualization of user requirements can be specified using RG method, as shown in Fig. 5.

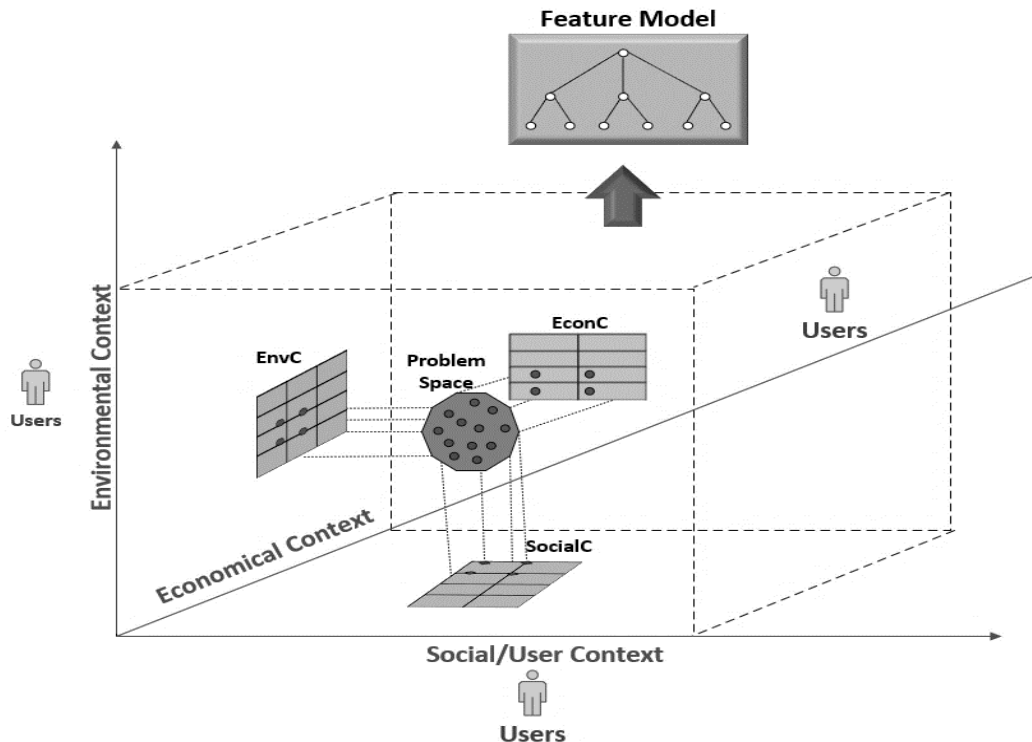


Fig.5 The conceptual scheme of RG method usage [11]

To continue our approach to formalization of the RG method, we can represent this space as a subset of the Cartesian product for the three sets defined in (1), namely:

$$PS(RG) \subseteq E \times C \times R \quad (4.2)$$

At the stage of SHS domain research stakeholders are user experts and the processes of knowledge elicitation, RG creation and complexity analyzation are accompanied by domain experts and knowledge engineers.

Given the fact that the above-mentioned aspects determine when and under what conditions one or another system function(s) should be used, RG can help to identify the impact of the selected function(s) on the behavior of the whole system, to reflect the commonality and variability between them.

From the interviews of domain experts in SHS design in terms of social context it is possible to elicit:

- elements: «Sleeping», «Not at home», «Leaving house», «Reduce the expenses» ...;
- constructs: «Temperature should be/doesn't be lower», «Soft music can be/needs not to be played», «Light should be turned on/off», «Windows should be opened/closed»...;
- rating Values: in range of [1...5].

Fig. 6 represents the created RG based on elicited elements, constructs and rating values from users' interviews.



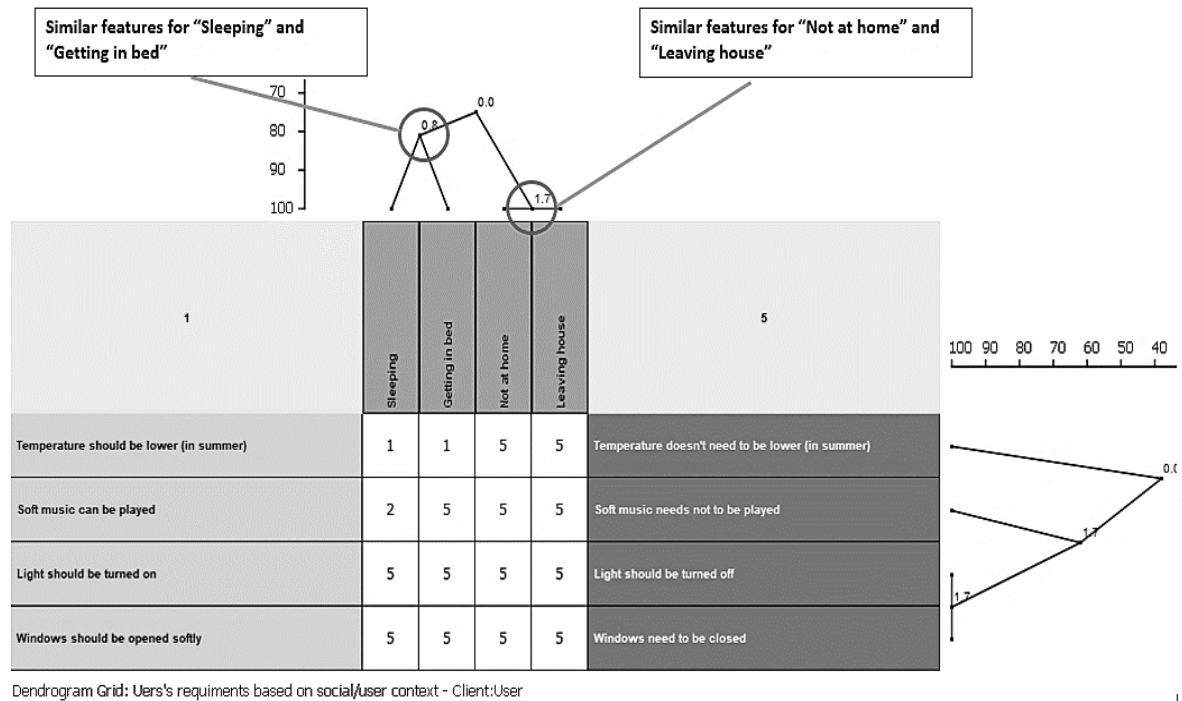


Fig.6 The structured knowledge analysis with usage of RG method

Through the elicited SHS properties and created RG dendrogram it is possible to analyze the common and variable points and reveal the complexity of the system. First, the estimations received from the interviews are analyzed. As shown on the Fig. 6, the construct “Light should be turned on/Light should be turned off” gets 5 for any context. This rating means that for all proposed contexts, function of “turned on light” is useless as it will never be used. The analysis of estimations for other elements occurs similarly. Thus, with RG dendrogram it is possible to determine which functions in the particular context are relevant and which are unnecessary.

Next step of RG dendrogram analysis is detection of common and variable aspects. Based on the given RG, common features will exist for “Sleeping” and “Getting to bed” as they both, due to the context of “Temperature should be lower/ Temperature doesn’t need to be lower”, “Light should be turned on/Light should be turned off” and “Windows should be opened softly/Windows need to be closed”, have the same rating values. Likewise, common features can be identified for “Not at home” and “Leaving house” as both of these contexts have the same rating value for all constructs. The variable point of this particular RG would be the function “Soft music can be played/Soft music needs not to be played” for the social context of “Sleeping”, as the exact functionality that should be used during this time is depends on the user attitude or other possible circumstances.

### 5. The proposed approach to building of a target FODA-model for variability management

To continue our research of the usage of RG method in handling the expert knowledge output for effective variability modeling in SPL development, we propose to consider the requirements specification in the form of a user story (UrS). The UrS is a special form which represents initial user requirements in semi-structured notation [12], where the following syntax is to be used for user needs specification in a natural language:

$$UrS \Rightarrow \{As\ a\ <\text{“User’s role”}\ >, I\ want\ <\text{“Action”}\ >\ so\ that\ <\text{“Goal”}\ >\}, \quad (5.1)$$

where the definitions “User’s role”, <“Action”>, and “Goal” are the appropriate textual descriptions (usually a simple sentences) given in terms of a target problem domain.

The UrS approach is a recognized way to exactly describe the domain specifics in agile-software development, where possible changes in user requirements must be taken into account. With respect to

our target problem domain: SHS development, some possible examples of UrS can be represented as follows:

UrS1: {As a “SHS Renter”, I want “Light to be turned off” so that “I can sleep/leave the house”},

UrS2: {As a “SHS Owner”, I want “Temperature to be lower” so that “I can cut the expenses”},

UrS3: {As a “SHS Installer” I want “Types of sensors to be less” so that “I can reduce time to configurate”},

.....

If now we compare the semantical meanings of the definitions {“User’s role”, <”Action”>, and ”Goal”} used in the general syntax form (3) with respect to their examples in *UrS1*, *UrS2*, *UrS3*..., with the components of RG represented in formula (1), it makes clear, that the definition {“User’s role”} corresponds to the elements of a set *E*, and the definitions {<”Action”>} and {”Goal”} match to the elements of a set *C* respectively. In the formal way it can be defined as two mapping rules  $\rho$  and  $\mu$  accordingly, and represented as:

$$\rho: E \rightarrow \{“User’s role”\}, \mu: C \rightarrow \{<“Action”>, “Goal”\} \tag{5.2}$$

It has to be mentioned that in recent publications about UrS technique the possibility of their usage for conceptual domain modeling is emphasized [12, 13], and there are already some CASE-tools to support a mapping of UrS to appropriate domain models, e.g. to FODA-models [1,4]. That is why it is possible to consider the sequential usage of the formal definitions for RG and UrS given in (4.1, 4.2) – (5.1, 5.2) as a description of a *structured information base* for the elaboration of a computerized procedure to construct a target FODA-model. The example of such a model for SHS development is shown in Fig. 7, as the structured description of technological variability features to be provided in the target system (in comparison with its initial vision given in Fig. 3).

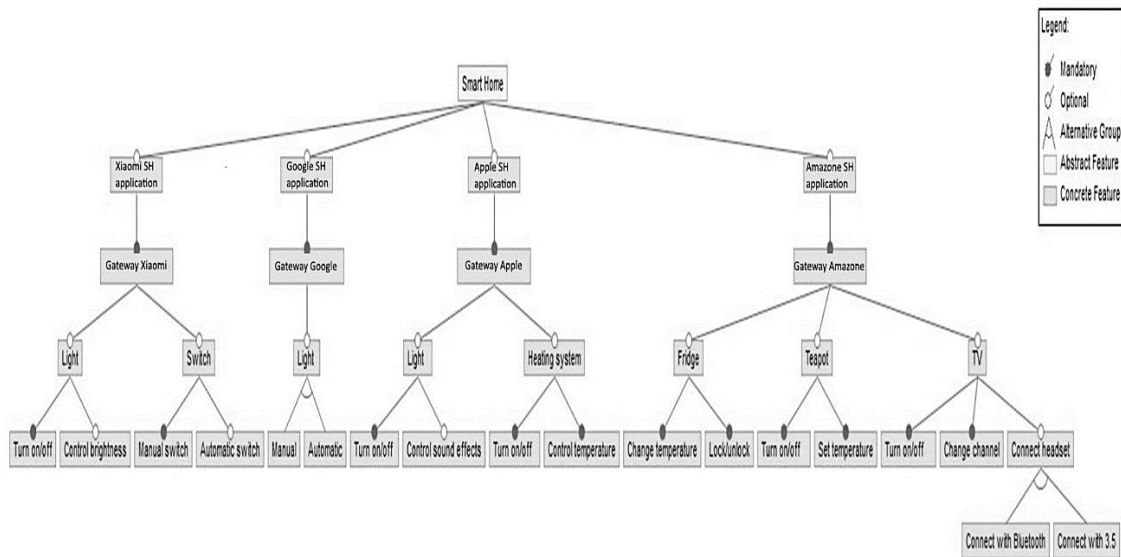


Fig.7 The example of a variability model in FODA-notation for SHS development

To implement an appropriate computerized procedure to automated creation of a target FODA-model using RG methods for expert knowledge handling it is also necessary to apply some CASE-tools for supporting this process [15], and the GridSuite is one of such tools. [16].

### 6. Conclusions and further work

In this paper the actual scientific and technical problem: the development of software product lines (SPL) with respect to variability issues at all stages of their full life cycle (FLC) is considered. One of the most important stages in FLC is the requirements engineering, and to support this process it is advisable to utilize the methods of expert knowledge elicitation and analysis. The review of some

methods to extract and to process of expert knowledge has been carried out, their possible classification has been offered, and the method of repertory grids (RG) has been chosen for the further usage. The essentials of the RG method are considered and its formal definition is elaborated. The examples in the domain of smart house systems using this approach is considered which allows analyzing the technical, social and economic aspects of user requirements for software and hardware variability at the conceptual design stage of these systems. As the result, a possibility to automate a process of a corresponding FODA-model creation that provides an automated variability support at the all FLC stages for SPL has been shown.

For the next step we are going to develop the appropriate information technology to support the proposed approach using some already available CASE-tools and our own software solutions to be elaborated.

#### REFERENCES

1. K. Pohl, et al: *Software Product Line Engineering: Foundations, Principles, and Techniques*. Springer, 2005, 467 pp.
2. E. Evans, *Domain-Driven Development (DDD): Structuring of complex software systems: translation from eng.* – M. Williams, 2011, 448pp. [in Russian]
3. K. M. Lavrishcheva, et al. *New theoretical bases of technology production of software systems families in the context of generating programming: monograph* / In-t program. systems. NAS Ukraine, - K., 2011, pp. 277. [in Ukrainian]
4. M. Tkachuk, R. Gamzaev, I. Martinkus et al. *Towards Effectiveness Assessment of Domain Modelling Methods and Tools in Software Product Lines Development* // Enterprise Modelling and Information Systems Architectures – International Journal of Conceptual Modeling, Vol. 13 (2018), Germany, pp. 190-206.
5. J. Tiihonen, M. Raatikainen, V. Myllärniemi, et al. *Carrying Ideas from Knowledge-based Configuration to Software Product Lines* // Proceedings of the 15th International Conference on Software Reuse, Cyprus, June 5-7, 2016, pp. 55–62.
6. Official Web-site of System Thinking World community, 2004. [Online]. Available: <http://www.systems-thinking.org/kmgmt/kmgmt.htm> (Last accessed on 12.09.2020).
7. Saeed F, et al. *IoT-based Intelligent Modeling of Smart - Home Environment* // Journal of Sensor and Actuator Networks. 2019; 7(1).
8. R.O. Gamzayev, M.V. Tkachuk, O.T. Tovstokorenko. *Application of domain modeling methods to support software variability in the development of “Smart-Home” systems* // “Information systems and technologies” ICT-2020. Materials of 9<sup>th</sup> International scientific and technical conference September 17-20, 2020, Kobleve – Kharkiv, Ukraine (included in the conference program) [in Ukrainian]
9. Methods of knowledge acquisition automatization in expert systems: classification, current state, comparative analysis // International scientific-practical journal “Software products and systems”. [Online]. Available: <http://www.swsys.ru/index.php?page=article&id=1187&lang=.docs>. (Last accessed on: 12.09.2020). [in Russian]
10. M. L. G. Shaw, B. R. Gains. *On the relationship between repertory grid and term subsumption knowledge structures: theory, practice and tools.* // Research and Development in Expert Systems, M.A.Bramer & R.W.Milne (Eds), CUP, 1993, pp. 125-143.
11. D. Sangeeta, L. Seok-Won. *From Requirements Elicitation to Variability Analysis Using Repertory Grid: A Cognitive Approach* // Proceeding of RE 2015, Ottawa, ON, Canada, 2015, pp. 46-55.
12. A. Gupta. *Generation of Multiple Conceptual Models from User Stories in Agile* // Joint Proceedings of REFSQ-2019 Workshops, Doctoral Symposium, Live Studies Track, and Poster Track, Essen, Germany, 18-03-2019, published at <http://ceur-ws.org>.
13. R. Mesquita, A. Jaqueira, C. Agra, et al. *US2StarTool: Generating i\* Models from User Stories* // Proceedings of the Eighth International i\* Workshop (ISTAR-2015), CEUR Vol-978.
14. G. Lucassen, F. Dalpiaz, J. Martijn, et al. *Improving Agile Requirements: the Quality User Story Framework and Tool.* In: *Requirements Engineering*, Vol. 21 (2016), pp. 383–403.
15. A. Curtis, T.M. Well, T. Higbee, et al. *An Overview and Tutorial of the Repertory Grid Technique in Information Systems Research.* In: *Communications of the Association for Information Systems (CAIS)*, vol. 23(3), pp. 37-62.
16. GridSuite: official We-site of the Repertory Grid Software. 2003-2020. [Online]. Available: <https://www.gridsuite.de/45623/41431.html> (Last accessed on: 12.09.2020).

#### ЛІТЕРАТУРА

1. Pohl K., et al: Software Product Line Engineering: Foundations, Principles, and Techniques., Springer, 2005. 467 pp.
2. Эванс Э. Предметно-ориентированное проектирование (Domain-Driven Development-DDD): Структуризация сложных программных систем: пер. с англ. - М. Вильямс, 2011. 448 с.
3. Лавріщева К. М. та ін. Нові теоретичні засади технології виробництва сімейств програмних систем у контексті генерувального програмування : монографія / Ін-т програм. систем. НАН України, - К., 2011. 277 с.
4. Mykola Tkachuk, Rustam Gamzaev, Iryna Martinkus et al. Towards Effectiveness Assessment of Domain Modelling Methods and Tools in Software Product Lines Development // Enterprise Modelling and Information Systems Architectures – International Journal of Conceptual Modeling, Vol. 13 (2018), Germany. pp. 190-206.
5. Tiihonen, J., Raatikainen, M., Myllärniemi, V. et al. Carrying Ideas from Knowledge-based Configuration to Software Product Lines // Proceedings of the 15th International Conference on Software Reuse, Cyprus, June 5-7, 2016, pp. 55–62.
6. Official Web-site of System Thinking World community, 2004. URL: <http://www.systems-thinking.org/kmgmt/kmgmt.htm> (дата звернення 12.09.2020).
7. Saeed F, et al. IoT-based Intelligent Modeling of Smart - Home Environment // Journal of Sensor and Actuator Networks. 2019; 7(1).
8. Гамзаєв Р.О., Ткачук М.В., Товстокоренко О.Т. Застосування методів доменного моделювання для підтримки варіабельності програмного забезпечення в розробці систем «розумний будинок» // «Інформаційні системи та технології» ІСТ-2020. МАТЕРІАЛИ 9-ї Міжнародної науково-технічної конференції 17-20 вересня 2020 р. Коблеве-Харків, Україна (включено до програми конференції).
9. Средства автоматизации приобретения знаний в экспертных системах: классификация, современное состояние, сравнительный анализ:[Электронный ресурс] // Международный научно-практический журнал «Программные продукты и системы». URL: <http://www.swsys.ru/index.php?page=article&id=1187&lang=.docs>. (дата звернення: 12.09.2020).
10. Shaw M.L.G., Gains B.R. On the relationship between repertory grid and term subsumption knowledge structures: theory, practice and tools. // Research and Development in Expert Systems, M.A.Bramer & R.W.Milne (Eds), CUP, 1993, pp. 125-143.
11. Sangeeta Dey, Seok-Won Lee. From Requirements Elicitation to Variability Analysis Using Repertory Grid: A Cognitive Approach // Proceeding of RE 2015, Ottawa, ON, Canada, 2015. pp. 46-55.
12. Gupta, A. Generation of Multiple Conceptual Models from User Stories in Agile // Joint Proceedings of REFSQ-2019 Workshops, Doctoral Symposium, Live Studies Track, and Poster Track, Essen, Germany, 18-03-2019, published at <http://ceur-ws.org>
13. Mesquita, R., Jaqueira, A., Agra, C. et al US2StarTool: Generating i\* Models from User Stories // Proceedings of the Eighth International i\* Workshop (ISTAR-2015), CEUR Vol-978.
14. Lucassen, G., Dalpiaz, F., Martijn J. et al. Improving Agile Requirements: the Quality User Story Framework and Tool. In: Requirements Engineering, Vol. 21 (2016), pp. 383–403.
15. Curtis, A., Wells, T.M., Higbee, T. et al. An Overview and Tutorial of the Repertory Grid Technique in Information Systems Research. In: Communications of the Association for Information Systems (CAIS), vol. 23(3), pp. 37-62.
16. GridSuite: official We-site of the Repertory Grid Software. 2003-2020. URL: <https://www.gridsuite.de/45623/41431.html> (дата звернення: 12.09.2020).

УДК 004.9

## Алгоритм побудови моделі веб-сайту

Н.А. Гук, С.В. Диханов, О.Д. Матющенко

**Гук  
Наталія Анатоліївна**

*д.ф.-м.н., професор; завідувач кафедри комп'ютерних технологій  
Дніпровський національний університет імені Олеся Гончара, Дніпро, пр.  
Гагаріна, 72, Дніпро, Україна, 49010  
e-mail: [natalygyuk29@gmail.com](mailto:natalygyuk29@gmail.com);  
<https://orcid.org/0000-0001-7937-1039>*

**Диханов  
Станіслав Віталійович**

*аспірант  
Дніпровський національний університет імені Олеся Гончара, Дніпро, пр.  
Гагаріна, 72, Дніпро, Україна, 49010  
e-mail: [dykhanovstas@gmail.com](mailto:dykhanovstas@gmail.com);  
<https://orcid.org/0000-0001-9073-0784>*

**Матющенко  
Олег Дмитрович**

*студент  
Національний технічний університет «Харківський політехнічний  
інститут», вул. Курпичова 2, Харків, Україна, 61002  
e-mail: [matyushchenkk@gmail.com](mailto:matyushchenkk@gmail.com);  
<https://orcid.org/0000-0003-3596-5737>*

У роботі для аналізу структури веб-сайту запропоновано використовувати модель у вигляді графа. Для побудови моделі у вигляді веб-графу розроблено метод, алгоритм сканування сторінок веб-ресурсу. За допомогою фреймворку Scrapy та мови Python побудовано програмне забезпечення та виконано сканування веб-ресурсів з метою побудови їх веб-графів. Для візуалізації отриманих графів та обчислення деяких метричних характеристик застосовано Gephi. Виконано сканування веб-сайтів, побудовано веб-графи, з використанням метричних характеристик здійснено аналіз структурної зв'язності.

**Ключові слова:** веб-сайт, веб-граф, краулінг, Scrapy, Gephi.

## Algorithm for building a website model

N. Huk, S. Dykhanov, O. Matiushchenko

**Huk Natalia**

*Doctor Sciences in Physics and Mathematics, Professor; Head of the  
Department of Computer's Science  
Oles Honchar Dnipro National University, 72, Gagarin Avenue, Dnipro,  
Ukraine, 49010*

**Dykhanov Stanislav**

*PhD student  
Oles Honchar Dnipro National University, 72, Gagarin Avenue, Dnipro,  
Ukraine, 49010*

**Matiushchenko Oleh**

*Student  
National Technical University «Kharkiv Polytechnic Institute» 2, Kyrpychova  
str., Kharkiv, Ukraine, 61002*

The analysis of the structure of the website modeling has been carried out. The models of Internet space representation in the form of semantic networks, frame structures and ontology have been analyzed. The web graph model has been chosen to represent the web resource. The pages of a web resource are connected by hyperlinks, which form the internal structure of the resource. To build a model of a website in the form of a web graph, a method and algorithm for scanning the pages of a web resource have been developed. The web resource scanning is performed by in depth searching with the LIFO (Last In - First Out) method. Links are searched by sorting the lines of the page markup text and extracting links by using regular expressions. Only links to pages within the resource are taken into account in the search process, external links are ignored. The crawling procedure is implemented by using the Scrapy framework and the Python. To account for the presence of additional filters used to select pages with criteria, the rules for selecting URL in HTML code have been strengthened. Web resources are scanned to build their web graphs. Storing information by using a list of edges and an adjacency matrix is used in further work with the obtained web graphs. To visualize the obtained graphs and calculate some metric characteristics, the Gephi software

environment and the algorithm for stacking the vertices of the Yifan Hu graph has been used. The graph diameters, the average vertex degree, the average path length, the density factor of the graph are used for analysis of the structural connectivity of the graphs studied. The proposed approach can be applied during the site reengineering procedure.

**Keywords:** website, web graph, crawling, Scrapy, Gephi.

## Алгоритм построения модели веб-сайта

Н.А. Гук, С.В. Дыханов, О.Д. Матющенко

<b>Гук Наталья Анатольевна</b>	<i>д.ф.-м.н., профессор; заведующая кафедрой компьютерных технологий Днепропетровский национальный университет имени Олеся Гончара, пр. Гагарина, 72, Днепро, Украина, 49010</i>
<b>Дыханов Станислав Витальевич</b>	<i>аспирант Днепропетровский национальный университет имени Олеся Гончара, пр. Гагарина, 72, Днепро, Украина, 49010</i>
<b>Матющенко Олег Дмитриевич</b>	<i>студент Национальный технический университет «Харьковский политехнический институт», ул. Кирпичева 2, Харьков, Украина 61002</i>

В работе для анализа структуры сайта предложено использовать модель в виде графа. Для построения модели в виде веб-графа разработан метод, алгоритм сканирования страниц веб-ресурса. С помощью фреймворка Scrapy и языка Python построено программное обеспечение и выполнено сканирование веб-ресурсов с целью построения их веб-графов. Для визуализации полученных графов и вычисления некоторых метрических характеристик используется Gephi. Выполнено сканирование веб-сайтов, построены веб-графы, с использованием метрических характеристик осуществлен анализ структурной связности.

**Ключевые слова:** веб-сайт, веб-граф, краулинг, Scrapy, Gephi

### 1 Вступ

Процес створення або реконструкції будь-якого веб-сайту починається з аналізу його структури. Грамотне проектування розділів, підрозділів і кінцевих сторінок робить сайт зручним та зрозумілим для користувача, а також покращує помітність ресурсу у мережі Інтернет для пошукових робіт, які формують видачу за запитами користувачів. Тому при розробці та реінжинірингу існуючих сайтів завжди намагаються приділити увагу логіці подачі матеріалу.

Якщо ресурс має великий обсяг та складну ієрархію розділів, і грамотно не структурований, то перебування на такому сайті і пошук інформації стають проблематичними для більшості користувачів, також ускладнюється і індексація ресурсу пошуковими роботами.

Для усунення можливих помилок в організації і логіці структури ресурсу, при проведенні внутрішнього та технічного аудитів виконується аналіз структури сайту.

Для проведення подібного аналізу використовуються моделі, за допомогою яких зображення складної структури стає зрозумілим.

Моделі веб-сайту можуть слугувати аналітичним інструментом, пояснювальним інструментом, а також інструментом для прогнозування подальшого розвитку ресурсу [1], тому питання розробки моделей та алгоритмів їх побудови є актуальною задачею.

### 2 Огляд літератури

На сьогодні існує декілька моделей, за допомогою яких можна зображувати інтернет-ресурси.

Онтології [2] є ефективним засобом представлення та систематизації всіх видів інформації про модель предметної області. Вони використовуються для формалізації, систематизації та збору знань і даних, надають можливість проводити обробку та аналіз отриманих структур, а також забезпечують організацію зручного змістовного доступу до об'єктів. За допомогою онтологій існує можливість надати умовні характеристики об'єкту (атрибути) для понять предметної області, встановлювати семантичні взаємозв'язки (відношення) між елементами, заповнювати побудовану онтологію екземплярами. Редактори онтологій та даних здійснюють процес налаштування та управління системою знань і контентом.

У роботі [3] наведено застосування онтологій для інтернет-технологій, побудовано моделі на мові OWL (Ontology Web Language) та показано, що подібні моделі можна застосовувати для представлення знань та управління ними. Також на основі онтологій можна створювати зручну навігацію по інтернет-ресурсам, адекватну інтерпретацію змісту текстових документів та пошукових запитів.

Однак побудова онтологій для веб-ресурсів є трудомістким процесом, який складно застосовувати для оперативного аналізу окремих сайтів.

За допомогою семантичної мережі існує можливість отримати модель веб-сайту в вигляді орієнтованого графу, вершинами якого є сторінки ресурсу, а дуги графа можуть відображати семантичні відношення між сторінками. Взагалі семантичні відношення розділяються на лінгвістичні, за допомогою яких встановлюється змістовний зв'язок між об'єктами, логічні відношення, за допомогою яких існує можливість обчислювати висловлювання, теоретико-множинні відношення, які встановлюють зв'язок типу «частина-ціле», та квантифіковані відношення, які використовують квантори узагальнення та існування.

Однак таку модель краще використовувати для інформаційного пошуку, побудованого з використанням семантичного значення змісту або інформації про сторінку.

Зображення веб-сайтів у вигляді фреймової моделі наведено у роботі [4], де модель сайту представлено у вигляді мережі фреймів, що містять інформацію про сторінки, набори їх властивостей та зв'язки між ними.

Використання подібної моделі є зручним для організації пошуку, порівняння сайтів між собою, виявлення однотипних сторінок та структур посилань.

Перелічені моделі найчастіше застосовуються, коли необхідно проаналізувати не лише структуру, а і змістовні зв'язки між сторінками, а також для набуття нових знань методами Web Data Mining [5].

Для оперативного аналізу структури сайтів та фрагментів веб-простору у якості моделей найчастіше використовуються веб-графи [6-10]. За допомогою такої моделі зручно відстежувати зв'язки між сторінками, аналізувати поведінку користувачів при переміщенні по сайту [11, 12], відокремлювати схожі за ознаками сегменти веб-простору.

Для побудови такої моделі веб-сайту шляхом отримання інформації про гіперпосилання сторінок у мережі Інтернет застосовується процедура веб-краулінгу. Основною задачею краулінгу є організований обхід веб-простору та збір інформації про гіперпосилання сторінок. Основні принципи роботи краулерів наведено у роботах [13-15].

Програми-краулери розділяють на універсальні та спеціалізовані. Універсальні здійснюють обхід веб-простору з використанням посилань, аналізують сторінки на наявність необхідної інформації, процес виконується до тих пір, поки не буде виконане обмеження на кількість проаналізованих сторінок або не досягнуто мети пошуку. Вони найчастіше є надлишковими та вимагають значної кількості обчислювальних ресурсів, програють у продуктивності спрямованого інформаційного пошуку за заданою метрикою значущості.

Спеціалізовані краулери розробляються для збору заздалегідь визначеного типу інформації, наприклад, для дослідження сайтів організацій, які працюють у одному сегменті ринку. В роботах [16-18] наведено огляд існуючих веб-краулерів з відкритим вихідним програмним кодом та безкоштовних розповсюджених сервісів для здійснення процедури краулінгу, інструментів для збору та аналізу даних з сайтів. Розглянуті інструменти класифіковано відповідно до необхідних для їх роботи ресурсів та обмежень щодо виконання спеціалізованих завдань.

В роботі [19] запропоновано модель веб-краулера, який складається з узагальненого ядра та спеціалізованих додатків, додатки налаштовуються для здійснення узагальненого пошуку по гіперпосиланням, тематичного пошуку з аналізом текстів сторінок, пошуку по популярним сторінкам та інші типами пошуку.

З аналізу літературних джерел стає зрозуміло, що для проведення аналітичних досліджень веб-простору та його сегментів можливе застосування результатів розширеного пошуку, здійсненого відомими пошуковими системами, використання відкритих баз даних, також можна здійснювати закупівлю вебметричних даних у аналітичних компаній, однак значну перевагу має розробка власних методів та алгоритмів збору даних про досліджувані об'єкти, оскільки це дозволяє налаштовувати інструменти збору даних під вимоги, які виникають під час проведення дослідження.

### **3 Постановка задачі**

Для побудови веб-графу необхідно розробити метод сканування сторінок сайту (процедуру краулінгу), алгоритм та відповідне програмне забезпечення, за допомогою якого можна здійснювати обхід та сканування веб-сайту, будувати веб-граф у вигляді списку ребер та матриці суміжності. Результат сканування необхідно представити у зручному вигляді для подальшої обробки та візуалізації.



#### 4 Математична модель веб-сайту

Гіпертекстову модель веб-сайту  $H$  можна представити як набір, що складається з двох множин:

$$H = \{P, L\}, \quad (4.1)$$

де  $P = \{p_1, \dots, p_n\}$  – множина сторінок сайту;  $L = \{l \mid \exists p_1, p_2 \in P: p_1, p_2 \in l(p_1, p_2)\}$  – множина гіперпосилань між сторінками.

Гіперпосилання є базовим елементом веб-простору, що зв'язує між собою веб-ресурси та веб-сторінки одного ресурсу. У найпростішому вигляді можна вважати, що гіперпосилання - це пара  $\langle \text{URL-start}, \text{URL-end} \rangle$ , де URL-start є адресою веб-ресурсу (або веб-сторінки), з якого зроблено посилання. Це реалізує можливість переходу на веб-ресурс (веб-сторінку) з адресою URL-end. Зокрема, гіперпосилання, що з'єднують сторінки і документи одного веб-ресурсу, задають його внутрішню структуру та є внутрішніми посиланнями. Саме зв'язки між сторінками одного веб-ресурсу розглядаються у роботі.

Структурі гіпертекстової моделі веб-сайту відповідає математична модель у вигляді орієнтованого незваженого графа  $G = (A, E)$ , у якому  $A = P$ ,  $E = L$ . У побудованому графі  $A$  – множина вершин, елементи якої описують сторінки сайту,  $E$  – множина ребер графу, елементи якої відповідають гіперпосиланням між сторінками.

#### 5 Метод та алгоритм побудови веб-графа

Процес збору даних о гіпертекстовій структурі ґрунтується на ідеях алгоритмів пошуку на графах з використанням локальної інформації. У структурі сторінки сайту необхідно знайти посилання на інші сторінки, зробити перехід по цим посиланням на інші сторінки, де знов необхідно здійснити пошук посилань.

Пошук посилань у веб-документі здійснюється шляхом перебору строк тексту розмітки сторінки та виділенні з них посилань за допомогою регулярних виразів, гіперпосилання вилучаються з html сторінки з тегів  $\langle a \rangle$ . У процесі пошуку враховуються лише посилання на сторінки всередині ресурсу, зовнішні посилання ігноруються. Знайдені посилання зберігаються у в файл або базу даних в вигляді записів сторінка  $A_i$  – сторінка  $A_j$  ( $\langle \text{URL-start}, \text{URL-end} \rangle$ ).

У веб-графі це відповідає такій структурі: вершина  $A_i$  – ребро  $e_{ij}$  – вершина  $A_j$ . Процес виконується циклічно доки не буде досягнута одна з умов: вичерпиться обмеження на глибину пошуку або буде знайдено всі посилання для заданого веб-ресурсу. Додатково до процедури пошуку висувуються такі вимоги:

- ігноруються посилання з протоколом передачі даних, відмінним від заданого;
- ігноруються результати запитів по гіперпосиланнях в тих випадках, коли відповідь сервера в http-заголовки не вказує Content-type text / html або не вказує Content-type зовсім, тобто об'єктами сканування є тільки html-сторінки, гіперпосилання, що вказують на файли з розширеннями .rar, .docx, .js та інші, не розглядаються;
- виконується нормалізація посилань для усунення випадків завантаження одній і тій же сторінки за посиланнями з різним написанням;
- сторінки з однаковим вмістом вважаються однією і тією ж сторінкою.

Для організації алгоритму побудови веб-графа застосовується структура даних стек або черга, до якої будуть потрапляти пов'язані одна з одною сторінки, а потім витягуватися сторінки веб-сайту для здійснення обходу. Відповідна структура даних обирається в залежності від мети побудови веб-графу, якщо використовується черга, то буде здійснюватися обхід веб-ресурсу в глибину, при використанні стеку організується обхід в ширину.

Запропонований підхід до побудови процедури краулінгу було реалізовано наступним алгоритмом.

Вхідні данні: посилання на головну сторінку веб-сайту, з якої починається формування веб-графу.

Вихідні данні: Результуючий файл с записами вершина  $A_i$ , вершина  $A_j$ , де  $A_i$  - адреса сторінки, на якій знайдено посилання на сторінку  $A_j$ .

Ініціалізація: організувати структуру даних  $T$  для реалізації алгоритму; організувати динамічний масив міток сторінок  $w[A_j]$ , які вже розглянуто алгоритмом.



Початок.

1. Обрати головну сторінку сайту  $A_1$  для початку обходу, розташувати  $A_1$  у структуру даних  $T$ , позначити сторінку, як таку, що пройдено,  $w[A_1] = 1$ .

2. Повторювати до тих пір, поки  $T \neq \emptyset$ :

Витягти сторінку  $A_i$  зі структури даних  $T$ , повернути її, як таку що пройдено.

Для всіх сторінок  $A_j$ , які пов'язані зі сторінкою  $A_i$  та відсутні у структурі даних  $T$ , виконати наступне:

Отримати посилання з HTML коду сторінки  $A_i$  на сторінку  $A_j$ , перевірити коректність адреси та існування сторінки  $A_j$ . Якщо умови виконано, то зберегти в результуючий файл запис: вершина  $A_i$ , вершина  $A_j$ . Додати сторінку  $A_j$  у структуру даних  $T$  та позначити її, як таку, що пройдено  $w[A_j] = 1$ .

Кінець.

Під час роботи алгоритму з HTML коду сторінки додатково можна отримувати інформацію про вміст сторінки (заголовок, елементи керування, рисунки), яка може бути збережена в якості властивостей сторінки для подальшого аналізу.

Побудований алгоритм краулінгу веб-ресурсу було реалізовано у вигляді прикладного програмного забезпечення, архітектуру якого наведено на рис 5.1.

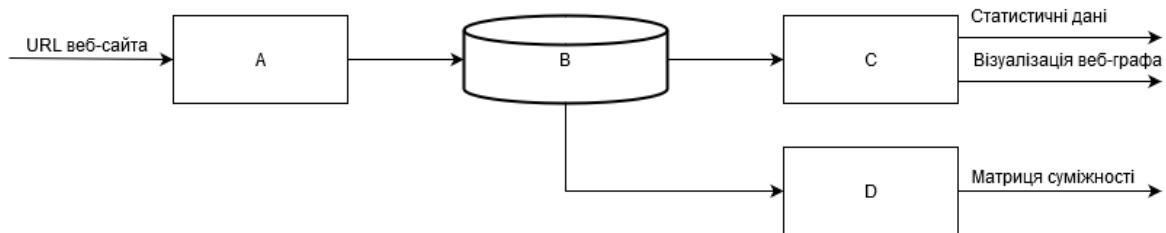


Рис. 5.1 Діаграма компонентів

Програмне забезпечення включає в себе чотири основних компоненти: програмний компонент А відповідає за побудову моделі веб-сайту та обробку даних. До його завдань входить сканування веб-сайту з метою відновлення гіпертекстової структури. Під час роботи відбувається завантаження веб-сторінок, витягування гіперпосилань з множини веб-сторінок, формування черги, нормалізація URL адрес, відсіювання веб-адрес з некоректними адресами, усунення дублювання адрес, формування списку суміжності у вигляді пар елементів <URL-start, URL-end> та для очищення цього списку суміжності від зайвих даних.

Компонент В забезпечує зберігання веб-графу сканованого веб-сайту у вигляді списку суміжності у форматі CSV.

Компонент С слугує для візуалізації графа, його укладання для поліпшення сприйняття та для обчислення деяких статистичних даних.

Компонент D здійснює перетворення списку суміжності в матрицю суміжності та пошук унікальних вершин.

## 6 Особливості програмної реалізації

Відповідно до розробленої архітектури програмна система побудована у вигляді консольних додатків з використанням мови програмування Python (компоненти А, D), а також із застосуванням існуючого у відкритому доступі програмного забезпечення Gephi для візуалізації побудованого графу (компонента С). Крім стандартних засобів платформи Java SE, використовувалися сторонні бібліотеки і компоненти.

Для здійснення сканування веб-сайту (компонента А) було використано вільно поширюваний Python-фреймворк Scrapy. Цей фреймворк забезпечує роботу по HTTP протоколу з будь-якою сторінкою, розташованою у мережі Інтернет. За його допомогою з HTML-коду сторінки зчитуються всі гіперпосилання на пов'язані сторінки, які потім зберігаються у відповідній

структурі даних. Дана операція повторюється для кожної сторінки, тим самим здійснюється обхід сайту. За отриманою гіпертекстовою структурою будується список суміжності.

Завантаження Scrapy може проводитись через PIP – систему управління пакетами, яка використовується для установки і управління програмними пакетами, що побудовано із використанням мови Python.

За замовчуванням Scrapy відвідує сторінки, використовуючи чергу LIFO (Last In – First Out), тобто використовує обхід веб-ресурсу пошуком в глибину.

Для здійснення процедури зберігання даних (Компонента В) було використано бібліотеку CSV, за допомогою якої компоненти матриці зберігаються в зручному для сприйняття користувачем і комп'ютером форматі. Цей формат має статичну структуру, тому для відслідковування помилок можна застосовувати автоматичний режим.

Для візуалізації побудованого графа використано GUI (Graphical user interface) додаток Gephi, який містить в собі набір основних способів укладань графів, а також інші інструменти для аналізу характеристик графів. В роботі застосовується алгоритм укладання Yifan Hu [20], що робить сприйняття отриманих результатів зрозумілим та наочним.

### 7 Аналіз результатів сканування веб-сайтів

Запропонована модель представлення веб-сайту, метод побудови краулеру та розроблене програмне забезпечення було застосовано для побудови та аналізу веб-графів існуючих у мережі Інтернет сайтів. Було розглянуто веб-сайти: офіційний сайт Дніпровського національного університету імені Олеся Гончара [Дніпровський національний університет імені Олеся Гончара [Електронний ресурс] – Режим доступу до ресурсу: <http://dnu.dp.ua/>], веб-сайт факультету прикладної математики Дніпровського національного університету імені Олеся Гончара [Сайт факультета прикладної математики ДНУ | [Електронний ресурс] – Режим доступу до ресурсу: <http://fpm.dnu.dp.ua/>], веб-сайт інтернет-магазину “Насіння країни” [Насіння країни [Електронний ресурс] – Режим доступу до ресурсу: <http://semena-dnepr.org.ua/>].

Обчислення виконувались із застосуванням апаратного забезпечення на двоядерному процесорі Intel® Core™ i5-5200U [60] частотою до 2.7 ГГц та 8 Гб оперативної пам'яті DDR3, відеокарті Nvidia GeForce 940M частотою до 1176 МГц та відеопам'яті 2 Гб DDR3.

У табл. 1 для порівняння наведено характеристики веб-сайтів (кількість сторінок та посилань між ними) та час сканування.

Таблиця 1

Параметри сайтів	Веб-сайти		
	dnu.dp.ua	fpm.dnu.dp.ua	semena-dnepr.org.ua
Кількість унікальних сторінок	24460	472	468
Кількість посилань між сторінками	2045729	26504	12096
Час сканування веб-ресурсу для побудови веб-графу, хв.	183	47	39

За результатами роботи розробленого програмного забезпечення були побудовані список суміжності та матриця суміжності веб-графів, на рис 7.1 наведено фрагмент списку суміжності для веб-графа сайту dnu.dp.ua.

<a href="http://www.dnu.dp.ua/view/kolisinini">http://www.dnu.dp.ua/view/kolisinini</a>	<a href="http://www.dnu.dp.ua/">http://www.dnu.dp.ua/</a>
<a href="http://www.dnu.dp.ua/news/3699">http://www.dnu.dp.ua/news/3699</a>	<a href="http://www.dnu.dp.ua/">http://www.dnu.dp.ua/</a>
<a href="http://www.dnu.dp.ua/news/3673">http://www.dnu.dp.ua/news/3673</a>	<a href="http://www.dnu.dp.ua/newsprint/3673">http://www.dnu.dp.ua/newsprint/3673</a>
<a href="http://www.dnu.dp.ua/news/3673">http://www.dnu.dp.ua/news/3673</a>	<a href="http://www.dnu.dp.ua/view/ffilol">http://www.dnu.dp.ua/view/ffilol</a>
<a href="http://www.dnu.dp.ua/news/3673">http://www.dnu.dp.ua/news/3673</a>	<a href="http://www.dnu.dp.ua/view/fsocgum">http://www.dnu.dp.ua/view/fsocgum</a>
<a href="http://dnu.dp.ua/view/our_partners">http://dnu.dp.ua/view/our_partners</a>	<a href="http://www.dnu.dp.ua/view/osvitni_programy">http://www.dnu.dp.ua/view/osvitni_programy</a>
<a href="http://dnu.dp.ua/view/our_partners">http://dnu.dp.ua/view/our_partners</a>	<a href="http://www.dnu.dp.ua/vybir_desciplin">http://www.dnu.dp.ua/vybir_desciplin</a>
<a href="http://dnu.dp.ua/view/our_partners">http://dnu.dp.ua/view/our_partners</a>	<a href="http://www.dnu.dp.ua/view/normativna_baza_oisvitnyogo_processu">http://www.dnu.dp.ua/view/normativna_baza_oisvitnyogo_processu</a>
<a href="http://www.dnu.dp.ua/map">http://www.dnu.dp.ua/map</a>	<a href="http://www.dnu.dp.ua/">http://www.dnu.dp.ua/</a>
<a href="http://www.dnu.dp.ua/view/legendu_dnu">http://www.dnu.dp.ua/view/legendu_dnu</a>	<a href="http://www.dnu.dp.ua/">http://www.dnu.dp.ua/</a>

Рис. 7.1 Фрагмент списку суміжності для сайту dnu.dp.ua.

Під час побудови веб-графу сайту інтернет-магазину було виявлено, що для товарів встановлено додаткові фільтри для вибору продукції за деякими критеріями, які знаходяться у тегу `<a href="">`. Якщо б краулер переходив по ним і записував усі знайдені ребра до списку суміжності, то зображення веб-графу було б некоректним, для кожної сторінки створювались би зайві посилання, а отриманий граф був би значно більшим. Для усунення зазначеної проблеми у правилах роботи Scrapy було посилено правила відбору URL в HTML-кодї, здійснювалось видалення ребер, у яких хоча б одна з URL-адрес (вершин графу) мала змінні.

Побудовані графи було зображено у графічному вигляді. За замовчуванням граф укладається випадковим чином – тобто розташування вузлів графу відносно друг друга відбувається довільно. Для спрощення сприйняття отриманих графів застосовуються різні укладки вершин. В роботі із застосуванням алгоритму Yuifan Hu вузли укладаються по колу у такий спосіб: вузли, які належать до однієї підмножини, притягуються один до одного.

На рис. 7.2 для порівняння впливу укладання на результуюче зображення графу досліджуваний граф сайту `dnu.dp.ua` зображено довільним чином та із застосуванням укладання. Слід зазначити, що застосування алгоритмів укладання значно спрощує сприйняття структури графу, можна бачити гетерогенність структури сайту.

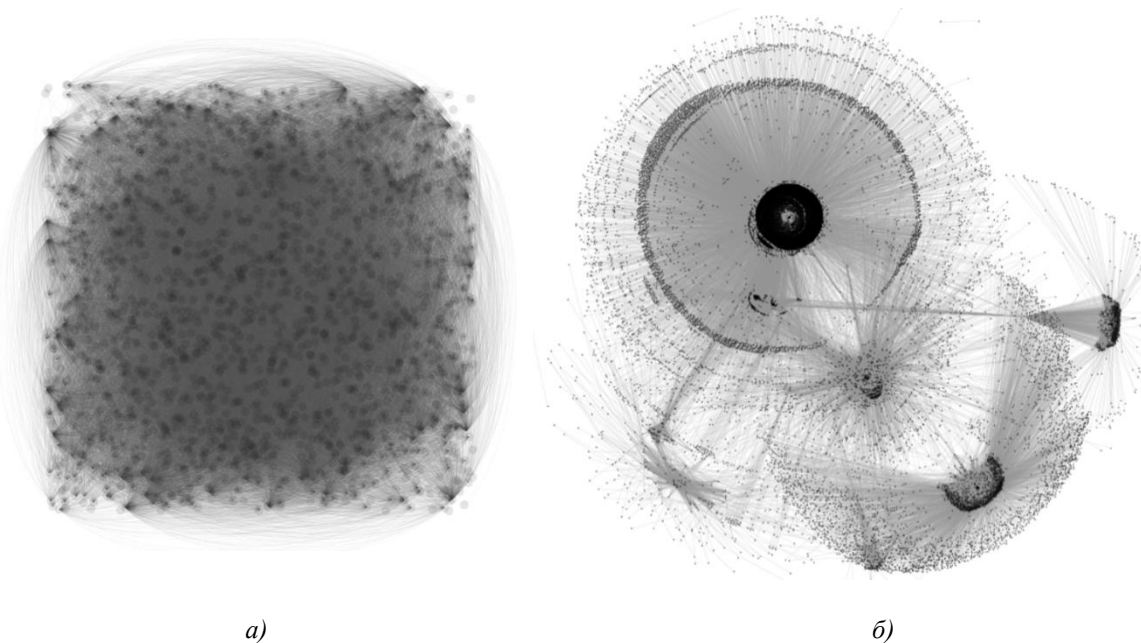


Рис. 7.2 а) граф `dnu.dp.ua` без укладання; б) граф `dnu.dp.ua` з укладанням Yuifan Hu

Для аналізу отриманих графів було застосовано вбудовані у Gephi алгоритми для розрахунку статистичних показників графа. Серед метричних характеристик використовується діаметр графу, середня степiнь вершини, середня довжина шляху, коефіцієнт щільності графа. Діаметр графа визначає максимальну відстань між двома вузлами через інші вузли. Середній ступiнь вузла дорiвнює середньої кількості вершин, з якими пов'язана вершина. Середня довжина шляху між двома вузлами дорiвнює середньої кількості ребер, що зв'язують вершини одну з одною.

Для аналізу зв'язності отриманої структури веб-сайту використовувалось такі показники, як коефіцієнт модулярності та коефіцієнт кластеризації графу. Коефіцієнт модулярності - це скалярна величина, що належить відрізьку  $[-1, 1]$  та зображує, наскільки при заданому розбитті графа на підмножини щільність зв'язків всередині підмножини більше, ніж щільність зв'язків між підмножинами, тобто визначає якість розбиття графу на підмножини. Коефіцієнт кластеризації зображує, наскільки сильно вершини схильні утворювати кластери (групи пов'язаних між собою вершин), які характеризуються тим, що вершини з однієї групи з'єднані між собою набагато щільніше, ніж з усім іншим графом.

Результати розрахунків зазначених величин наведено у таблиці 2:

Таблиця 2

Метричні характеристики	Веб-сайти		
	dnu.dp.ua	fpm.dnu.dp.ua	semena-dnepr.org.ua
Діаметр графа	207	5	13
Середня степінь вершин	60.561	18.845	18.2
Середня довжина шляху	21.039	3.042	2.306
Щільність графа	0.002	0.012	0.039
Коефіцієнт модулярності	0.206	0.189	0.11
Коефіцієнт кластеризації	0.33	0.24	0.176

Аналіз значень коефіцієнтів модулярності показав, що для графу сайту ДНУ цей показник значно вищий, ніж для інших досліджуваних сайтів. Високі показники коефіцієнту модулярності говорять про високу гетерогенність структури сайту, оскільки зростає кількість зв'язків всередині підграфів відносно до кількості зв'язків з іншими підграфами. Підтвердження цього висновку можна також отримати, якщо подивіться на рис. 7.2 з зображенням веб-графу сайту ДНУ з укладанням. Спостерігається наявність чіткої структури з можливістю бачити окремі зв'язні підграфи.

Наявність чіткої структури та існування зв'язних підграфів для сайту підтверджує й значення коефіцієнту кластеризації. Для сайту ДНУ цей показник дорівнює 0,33.

Сайти fpm.dnu.dp.ua та semena-dnepr.org.ua мають менший коефіцієнт модулярності та візуально не мають чітко виділених зв'язних підграфів, це підтверджує й більш низький коефіцієнт кластеризації, значення якого дорівнює 0,189 та 0,1 відповідно.

За результатами метричного аналізу можна робити висновки щодо зв'язності веб-графів та їх підграфів, відокремлювати тематичні кластери, перевіряти окремі веб-сторінки (наприклад, сторінки товарів інтернет-магазину) на належність до певних категорій, будувати рекомендаційні системи, виконувати процедури реінжинірингу сайтів.

## 8 Висновки

У роботі зроблено аналіз моделей представлення інтернет простору у вигляді семантичних мереж, фреймових структур, онтологій та обрано модель представлення веб-ресурсу у вигляді веб-графу. Для побудови моделі веб-сайту у вигляді веб-графу розроблено метод, алгоритм сканування сторінок веб-ресурсу та відповідне програмне забезпечення, що реалізує процедуру краулінгу. Програмне забезпечення побудовано із використанням фреймворку Scrapy та мови програмування Python, за допомогою розробленого програмного забезпечення можливе сканування веб-ресурсів з метою побудови їх веб-графів. Для подальшої роботи з отриманими веб-графами передбачено зберігання інформації у зручному вигляді за допомогою списку ребер та матриці суміжності. Для візуалізації отриманих графів та обчислення деяких метричних характеристик застосовано програмне середовище Gephi. Виконано сканування існуючих у мережі Інтернет веб-сайтів, побудовано їх веб-графи, із застосуванням метричних характеристик здійснено аналіз структурної зв'язності досліджуваних графів.

## ЛІТЕРАТУРА

1. Kumar R., Raghavan P., Rajagopalan S., Sivakumar D., Tomkins A., Upfal E. The Web as a graph. *In Proceedings of the 19th Symposium on Principles of Database Systems (PODS)*. ACM Press. 2000. P. 1-10. [https://www.researchgate.net/publication/221559387\\_The\\_Web\\_as\\_a\\_Graph](https://www.researchgate.net/publication/221559387_The_Web_as_a_Graph)
2. Ling Liu and Özsu Tamer M. Ontology to appear in the Encyclopedia of Database Systems [Електронний ресурс]. Springer-Verlag. 2008. URL: <http://tomgruber.org/writing/ontology-in-encyclopedia-of-dbs.pdf> (дата звернення: 06.10.2020)
3. Каунг М'ят Хту Анализ языка Веб онтологии (owl) и семантическая веб-технология. *Auditorium*. 2017. №4 (16). URL: <https://cyberleninka.ru/article/n/analiz-yazyka-veb-ontologii-owl-i-semanticheskaya-veb-tehnologiya> (дата звернення: 06.10.2020)

4. Ольшевский А.И., Кондратьева А.А. Описание способов представления web-сайтов в виде фреймовой модели для реализации функциональных операций в Интернет-клиентских системах. *Искусственный интеллект*. 2008. №1. С. 110–116. <http://dspace.nbuu.gov.ua/handle/123456789/6551>
5. Bing L. *Web Data Mining*. Springer. 2011. 624 p. <https://www.springer.com/gp/book/9783642194597>
6. Sheu P., Yu H., Ramamoorthy C., Joshi A., Zadeh L. Link Analysis in Web Mining: Techniques and Applications. *Semantic Computing Digital Object Identifier*. 2010. P. 69–86. [https://www.researchgate.net/publication/229459726\\_Link\\_Analysis\\_in\\_Web\\_Mining\\_Techniques\\_and\\_Applications](https://www.researchgate.net/publication/229459726_Link_Analysis_in_Web_Mining_Techniques_and_Applications)
7. Lai Wei, Huang Xiaodi From graph data extraction to graph layout: Web information visualization. *Information Sciences and Interaction Sciences (ICIS), 2010 3rd International Conference on Digital Object Identifier*. 2010. P. 224–229. <https://bit.ly/3omK2Qa>
8. Hall W., Tiropanis T. Web evolution and Web Science. *Computer Networks*. 2012. Vol. 56. № 18. P. 3859–3865. <https://bit.ly/3qB6NBG>
9. Шокин Ю.И., Веснин А.Ю., Добрынин А.А., Клименко О.А., Рычкова Е.В. Анализ веб-пространства академических сообществ методами вебометрики и теории графов. *Информационные технологии*. 2014. № 12. С. 31–40. <http://www.ict.nsc.ru/jspui/handle/ICT/252>
10. Ермолин Н.А., Мазалов В.В., Печников А.А. Теоретико-игровые методы нахождения сообществ в академическом Вебе. Труды СПИИРАН. 2017. 6(55), 237–254. URL: <https://doi.org/10.15622/sp.55.10> (дата звернення: 06.10.2020)
11. Горбунов А.Л. Марковские модели посещаемости веб-сайтов. *Интернет-математика 2007: сб. работ участников конкурса научных проектов по информационному поиску*. 2007. С. 65–73. <https://elar.urfu.ru/handle/10995/1334>
12. Liu Y., Ma Z. M., Zhou C. Web Markov Skeleton Processes and Their Applications // *Tohoku Mathematical Journal*. 2011. No. 63. P. 665–695. <https://bit.ly/3gg5vra>
13. Yadav M., Goyal N. Comparison of Open Source Crawlers- A Review. *International Journal of Scientific & Engineering Research*. 2015. V. 6. P. 1544-1551. <https://www.ijser.org/researchpaper/Comparison-of-Open-Source-Crawlers--A-Review.pdf>
14. Udupure T.V., Kale R.D., Dha R.C. Study of Web Crawler and its Different Types. *Journal of Computer Engineering*. 2014. Vols. 16(1). P. 3–4. <https://bit.ly/3qtBXep>
15. Ahuja M.S., Bal J.S., Varnica Web Crawler: Extracting the Web Data. *International Journal of Computer Trends and Technology*. 2014. №13(3). С. 132–137. <https://bit.ly/2VLLrrV>
16. Печников А.А., Сотенко Е.М. Программы-краулеры для сбора данных о представительских сайтах заданной предметной области – аналитический обзор. *Современные наукоемкие технологии*. 2017. № 2. С. 58-62. <https://top-technologies.ru/ru/article/view?id=36585>
17. Пудикова Е.М. Обзор веб-краулеров для решения задачи сбора данных о представительских сайтах заданной предметной области. *Системный анализ*. 2016. С. 1–16. <https://bit.ly/36PsTbL>
18. Гудков К.В., Тонкушин М.В. Анализ автоматизированных систем сбора информации в сети интернет. *Современные информационные технологии*. 2015. № 28. С. 27–31. [http://www.penzgtu.ru/fileadmin/filemounts/confcit/articles/autumn\\_2018/04.pdf](http://www.penzgtu.ru/fileadmin/filemounts/confcit/articles/autumn_2018/04.pdf)
19. Блеканов И.С., Сергеев С.Л., Мартыненко И.А. Построение тематико-ориентированных веб-краулеров с использованием обобщенного ядра. *Научно-технические ведомости санкт-петербургского государственного политехнического университета. Информатика. Телекоммуникации. Управление*. 2012. № 5(157). С. 9–15.
20. Hu Y. Efficient and high quality force-directed graph drawing. *Mathematica Journal*. 2006. №10. P. 37–71. <https://bit.ly/37W0bph>

## REFERENCES

1. R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, E. Upfal The Web as a graph. *In Proceedings of the 19th Symposium on Principles of Database Systems (PODS)*. ACM Press. 2000, P. 1–10. [https://www.researchgate.net/publication/221559387\\_The\\_Web\\_as\\_a\\_Graph](https://www.researchgate.net/publication/221559387_The_Web_as_a_Graph)
2. Ling Liu and Özsu Tamer M. Ontology to appear in the Encyclopedia of Database Systems [Электронный ресурс]. Springer-Verlag, 2008. URL: <http://tomgruber.org/writing/ontology-in-encyclopedia-of-dbs.pdf> (Last accessed: 06.10.2020)



3. Kaung Myat Htu “Web Ontology Language Analysis (owl) and Semantic Web Technology”, *Auditorium*, №4 (16), 2017. URL: <https://cyberleninka.ru/article/n/analiz-yazyka-veb-ontologii-owl-i-semanticheskaya-veb-tehnologiya> (date of access: 06.10.2020) [in Russian]
4. A.I. Olshevsky, A.A. Kondratyeva “Description of ways to represent web sites in the form of a frame model for the implementation of functional operations in Internet client systems”, *Artificial Intelligence*, №1, С. 110–116, 2008. <http://dspace.nbuv.gov.ua/handle/123456789/6551> [in Russian]
5. L. Bing *Web Data Mining*. Springer. 2011, 624 p. <https://www.springer.com/gp/book/9783642194597>
6. P. Sheu, H. Yu, C. Ramamoorthy, A. Joshi, L. Zadeh “Link Analysis in Web Mining: Techniques and Applications”, *Semantic Computing Digital Object Identifier*, P. 69–86, 2010. [https://www.researchgate.net/publication/229459726\\_Link\\_Analysis\\_in\\_Web\\_Mining\\_Techniques\\_and\\_Applications](https://www.researchgate.net/publication/229459726_Link_Analysis_in_Web_Mining_Techniques_and_Applications)
7. Lai Wei, Huang Xiaodi “From graph data extraction to graph layout: Web information visualization”, *Information Sciences and Interaction Sciences (ICIS), 2010 3rd International Conference on Digital Object Identifier*, P. 224–229, 2010. <https://bit.ly/3omK2Qa>
8. W. Hall, T. Tiropanis “Web evolution and Web Science”, *Computer Networks*, Vol. 56, № 18, P. 3859–3865, 2012. <https://bit.ly/3qB6NNG>
9. Yu.I. Shokin, A.Yu. Vesnin, A.A. Dobrynin, O.A. Klimenko, E.V. Rychkova “Analysis of the web space of academic communities using webometrics and graph theory”, *Information technologists*, № 12, С. 31–40, 2014. <http://www.ict.nsc.ru/jspui/handle/ICT/252> [in Russian]
10. N.A. Ermolin, V.V. Mazalov, A.A. Pechnikov “Game-theoretic methods for finding communities in the academic Web”, *SPIIRAS Proceedings*, 6(55), P. 237–254, 2017. URL: <https://doi.org/10.15622/sp.55.10> (date of access: 06.10.2020) [in Russian]
11. A.L. Gorbunov “Markov Models of Website Traffic”, *Internet Mathematics 2007: collection of articles. works of participants in the competition of scientific projects on information retrieval*. С. 65–73. 2007 <https://elar.urfu.ru/handle/10995/1334>
12. Y. Liu, Z.M. Ma, C. Zhou “Web Markov Skeleton Processes and Their Applications”, *Tohoku Mathematical Journal*. № 63. P. 665–695. 2011 <https://bit.ly/3gg5vra>
13. M. Yadav, N. Goyal “Comparison of Open Source Crawlers- A Review”, *International Journal of Scientific & Engineering Research*, V. 6, P. 1544-1551, 2015. <https://www.ijser.org/researchpaper/Comparison-of-Open-Source-Crawlers--A-Review.pdf>
14. T.V. Udupure, R.D. Kale, R.C. Dha “Study of Web Crawler and its Different Types”, *Journal of Computer Engineering*, Vol. 16(1), P. 3–4, 2014. <https://bit.ly/3qtBXep>
15. M.S. Ahuja, J.S. Bal, Varnica “Web Crawler: Extracting the Web Data”, *International Journal of Computer Trends and Technology*, №13(3), P. 132–137, 2014. <https://bit.ly/2VLlrV>
16. A.A. Pechnikov, E.M. Sotenko “Crawler programs for collecting data on representative sites of a given subject area - an analytical review”, *Modern high technology technologists*, № 2, С. 58-62, 2017. <https://top-technologies.ru/ru/article/view?id=36585> [in Russian]
17. E.M. Pudikova “Review of web crawlers for solving the problem of collecting data on representative sites of a given subject area”, *System analysis*, P. 1–16, 2016. <https://bit.ly/36PsTbL> [in Russian]
18. K.V. Gudkov, M.V. Tonkushin “Analysis of automated systems for collecting information on the Internet”, *Modern information technologies*, № 28, С. 27–31, 2015. [http://www.penzgtu.ru/fileadmin/filemounts/confcit/articles/autumn\\_2018/04.pdf](http://www.penzgtu.ru/fileadmin/filemounts/confcit/articles/autumn_2018/04.pdf) [in Russian]
19. I.S. Blekanov, S.L. Sergeev, I.A. Martynenko “Build subject-oriented web crawlers using a generic kernel”, *Scientific and technical statements of the St. Petersburg State Polytechnic University. Informatics. Telecommunications. Control*. № 5(157), С. 9–15, 2012. [in Russian]
20. Y. Hu “Efficient and high quality force-directed graph drawing”, *Mathematica Journal*, №10, P. 37–71, 2006. <https://bit.ly/37W0bph>

УДК 004.056 : 004.932

## Аналіз схеми автентифікації на основі використання QR-коду та веб-камери для пристроїв Smart-Mobile

О.І. Кот, І.І. Сватовський

**Кот Олександр Ігорович** *студент*  
*Харківський Національний Університет ім. В.Н. Каразіна*  
*майдан Свободи 4, 61022, Харків*  
*e-mail: xa10995161@student.karazin.ua ;*  
*https://orcid.org/0000-0002-6924-7445*

**Сватовський Ігор Іванович** *к.т.н., доцент*  
*Харківський Національний Університет ім. В.Н. Каразіна*  
*майдан Свободи 4, 61022, Харків*  
*e-mail: i.svatowsky@karazin.ua;*  
*https://orcid.org/0000-0002-1836-5599*

Проведено аналіз необхідності та доцільності використання методу автентифікації користувачів на основі QR-коду та веб-камери для пристроїв Smart-Mobile. Фішингові атаки є однією з найсерйозніших загроз, з якими стикаються користувачі в Інтернеті. Існуючі відомі схеми автентифікації в повній мірі не справляються з цими атаками, про що свідчать актуальні статистичні дані багатьох компаній, що досліджують питання кібербезпеки. В роботі пропонується нова схема автентифікації користувачів, яка дозволяє їм увійти в свої облікові записи без запам'ятовування паролів або наявності інших токенів автентифікації. Згідно протоколу обміну повідомленнями в запропонованій схемі користувач повинен відсканувати динамічно згенерований QR-код за допомогою програми для смартфона, а потім зробити власний фотознімок через веб-камеру, і відправити його на смартфон за допомогою повідомлення від сервера. Таким чином, повна процедура автентифікації вимагає мінімальної участі користувача і виконується в автоматичному режимі. Результати оцінки і практичного тестування показують, що запропонована схема автентифікації працює досить надійно і може використовуватися в якості безпечної схеми автентифікації користувачів для пристроїв Smart-Mobile.

**Ключові слова:** qr-код, веб-камера, аутентифікація, smart-mobile, Android.

## Analysis of the authentication scheme based on the use of QR-code and webcam for Smart-Mobile devices

O.I. Kot, I.I. Svatovskiy

**Oleksandr Kot** *Student*  
*V. N. Karazin Kharkiv National University*  
*4 Svobody Sq., Kharkiv, 61022, Ukraine*  
*e-mail: xa10995161@student.karazin.ua ;*  
*https://orcid.org/0000-0002-6924-7445*

**Igor Svatovskiy** *Ph.D., Associate Professor*  
*V. N. Karazin Kharkiv National University*  
*4 Svobody Sq., Kharkiv, 61022, Ukraine*  
*e-mail: i.svatowsky@karazin.ua;*  
*https://orcid.org/0000-0002-1836-5599*

The paper analyzes the necessity and expediency of using the method of user authentication based on QR-code and webcam for Smart-Mobile devices. Phishing attacks are one of the most serious threats faced by Internet users. Existing authentication schemes are not able to provide an adequate protection from these attacks, as evidenced by statistics collected by the companies researching cybersecurity. Therefore, the task of developing a secure authentication scheme for users, which can effectively counteract various types of phishing attacks is very important. The paper proposes a new authentication scheme for users, which allows them to log in to their accounts without remembering passwords or presenting other authentication tokens. According to the messaging protocol in the proposed scheme, the user must scan the dynamically generated QR-code using a smartphone application, then take their own photo via the webcam, and send it to the smartphone via a message from the server. Thus, the full authentication procedure requires minimal user involvement and is performed automatically. The results of evaluation and practical testing show that the proposed authentication scheme is quite reliable and can be used as a secure user authentication

scheme for Smart-Mobile devices. The proposed authentication protocol is not only able to cope with attacks such as Real Time Man-In-The-Middle and Controlled Relay Man-In-The-Middle, but can also protect users from the effects of malicious browser extensions and substitution of authentic applications by malicious variants. In addition, the proposed scheme does not require users to have any authentication tokens or credentials, as all they need is to scan the QR-code and verify the image taken by their own webcam. That makes the use of the proposed scheme more convenient and easy for users as compared to other known authentication schemes. Currently, the application of the proposed scheme requires the use of HTTPS websites for the exchange of all data involved. Thus, the proposed protocol can be implemented to manage cookies securely in order to prevent the interception of session data.

**Keywords:** *qr-code, web-camera, authentication, smart-mobile, Android.*

## Анализ схемы аутентификации на основе использования QR-кода и веб-камеры для устройств Smart-Mobile

А.И. Кот, И.И. Сватовский

**Кот Александр Игоревич**

*студент*

*Харьковский Национальный Университет им. В.Н. Каразина*

*площадь Свободы 4, 61022, Харьков*

*e-mail: xa10995161@student.karazin.ua ;*

*https://orcid.org/0000-0002-6924-7445*

**Сватовский Игорь  
Иванович**

*к.т.н., доцент*

*Харьковский Национальный Университет им. В.Н. Каразина*

*площадь Свободы 4, 61022, Харьков*

*e-mail: i.svatowsky@karazin.ua;*

*https://orcid.org/0000-0002-1836-5599*

Проведен анализ необходимости и целесообразности использования метода аутентификации пользователей на основе QR-кода и веб-камеры для устройств Smart-Mobile. Фишинговые атаки являются одной из самых серьезных угроз, с которыми сталкиваются пользователи в Интернете. Существующие известные схемы аутентификации в полной мере не справляются с этими атаками, о чем свидетельствуют актуальные статистические данные многих компаний, исследующих вопросы кибербезопасности. В работе предлагается новая схема аутентификации пользователей, которая позволяет им войти в свои учетные записи без запоминания паролей или наличия других токенов аутентификации. Согласно протокола обмена сообщениями в предложенной схеме пользователь должен отсканировать динамично сгенерированный QR-код с помощью приложения для смартфона, а затем сделать собственный фотоснимок через веб-камеру, и отправить его на смартфон с помощью сообщения от сервера. Таким образом, полная процедура аутентификации требует минимального участия пользователя и выполняется в автоматическом режиме. Результаты оценки и практического тестирования показывают, что предложенная схема аутентификации работает достаточно надежно и может использоваться в качестве безопасной схемы аутентификации пользователей для устройств Smart-Mobile.

**Ключевые слова:** *qr-код, веб-камера, аутентификация, smart-mobile, Android.*

### 1 Вступ

Фішингові атаки є однією з найсерйозніших загроз, з якими стикаються користувачі в Інтернеті: зловмисники намагаються вкрати конфіденційну інформацію, таку як дані для входу, дані кредитної картки, тощо, обманюючи користувачів для введення конфіденційної інформації на фішингових веб-сайтах і, таким чином, приводячи до величезних фінансових втрат [1]. Було запропоновано безліч схем виявлення фішингових атак, але кількість таких атак не зменшилася. З'явилися нові різновиди фішингових атак типу Active Man-In-The-Middle (MITM), які включають фішингові атаки Real Time Man-In-The-Middle (RT MITM) і Controlled Relay Man-In-The-Middle (CR MITM). Ці атаки дозволяють зловмисникам отримувати дані облікових записів користувачів і передавати їх в режимі реального часу. Так само зловмисник може спокусити користувача ввести дані у підробленому додатку і таким чином отримати доступ до облікового запису користувача [2-6]. Існуючі відомі схеми автентифікації в повній мірі не справляються з цими атаками. В роботі пропонується нова схема автентифікації користувачів, яка дозволяє їм увійти в свої облікові записи без запам'ятовування паролів або наявності інших токенів автентифікації. Згідно



протоколу обміну повідомленнями в запропонованій схемі користувач повинен відсканувати динамічно згенерований QR-код за допомогою програми для смартфона, а потім зробити власний фотознімок через веб-камеру, і відправити його на смартфон за допомогою повідомлення від сервера. Таким чином, повна процедура автентифікації вимагає мінімальної участі користувача і виконується в автоматичному режимі. Запропонована схема була реалізована і оцінена з точки зору зручності використання, можливості розгортання і параметрів безпеки. Результати оцінки і практичного тестування показують, що запропонована схема автентифікації працює досить надійно і може використовуватися в якості безпечної схеми автентифікації користувачів для пристроїв Smart-Mobile.

## **2 Запропонована схема автентифікації**

Пропонована схема багатофакторної автентифікації використовує довірений мобільний додаток і веб-камеру на клієнтському комп'ютері (настільному / портативному). Загальна процедура автентифікації користувачів за пропонованою схемою передбачає, що кожен раз, коли користувач хоче отримати доступ до свого облікового запису на веб-сайті, сервер відображає QR-код на веб-сторінці. Цей протокол автентифікації передбачає, що мобільний додаток в телефоні користувача є довіреним. Користувач сканує QR-код, що відображається на веб-сторінці, за допомогою мобільного додатку. Потім додаток відправляє дані користувача разом з токеном сеансу, отриманим з QR-коду, на адресу сервера. Далі сервер видає запит до клієнтського комп'ютера для доступу до веб-камери і робить знімок за її допомогою. Наступним етапом є відправлення зображення сервером в мобільний додаток за допомогою повідомлення. Отримавши це зображення, користувач підтверджує або відхиляє його. Після підтвердження правильності одержаного зображення користувач буде автентифікований. Таким чином, автентифікація сервера виконується за допомогою зображення, зробленого на клієнтській машині легітимного користувача, оскільки тільки законний веб-сайт може зробити знімок користувача за допомогою веб-камери і відправити вірне зображення на його мобільний телефон.

### **2.1 Припущення щодо необхідних умов для застосування**

1. Передбачається, що на ПК є веб-камера.
2. Передбачається, що користувач використовує справжній додаток Android під час реєстрації, і, як і в багатьох інших відомих схемах, процедура реєстрації нового користувача захищена від атак.
3. Запропонована схема передбачає, що передача даних між клієнтом і сервером відбувається за протоколом HTTPS, захищена від перехоплення в мережі і може використовуватися для обміну секретними ключами.
4. Автентичні сервери веб-сайтів та інформація, що зберігається в їх базах даних, вважаються безпечними.

### **2.2 Аналіз моделі загроз**

1. Фішинг, MITM-фішинг: зловмисник може заманити користувача на фішинговий веб-сайт, а потім отримати особисту інформацію. Зловмисник може або передати цю інформацію в режимі реального часу (RT MITM), або встановити модулі захоплення віддаленого робочого столу, або ретрансляції на термінал користувача (CR MITM) для крадіжки облікових даних.
2. Фішингові атаки на основі шкідливих розширень браузера: зловмисник може вкрасти облікові дані користувача, змусивши користувача встановити шкідливе розширення браузера, запитуючи дозволу на виконання деяких обережних дій у фоновому режимі, одночасно надаючи функції на передньому плані (front-end). Шкідливі розширення браузера можуть виконувати кейлоггінг, реєстрацію екрану або перехоплення пароля у фоновому режимі.
3. Спуфінг програми: зловмисник може створити підробний (фальшивий, підроблений) додаток для Android, який схожий на справжній додаток, необхідний для входу в систему. Потім зловмисник може встановити цей підроблений додаток на комп'ютер користувача і спокусити користувача ввести свої облікові дані через нього.

### **2.3 Процедура реєстрації користувачів**

Реєстрація користувача передбачає реєстрацію користувача в мобільному додатку. Таким чином, на етапі реєстрації беруть участь два об'єкти - мобільний додаток і веб-сервер. Реєстрація

користувача виконується в мобільному додатку, в якому будуть зберігатися дані користувача, які будуть використовуватися при вході в систему. Реєстрація користувача в запропонованому протоколі автентифікації показана на рисунку 2.1.

Алгоритм реєстрації користувача передбачає наступні кроки:

1. Користувач спочатку вводить відповідні дані, такі як ім'я користувача (UID), пароль (PWD), адреса електронної пошти (Email-ID), в мобільний додаток.
2. Мобільний додаток генерує випадкову сіль (salt), яка складається з 6 символів.

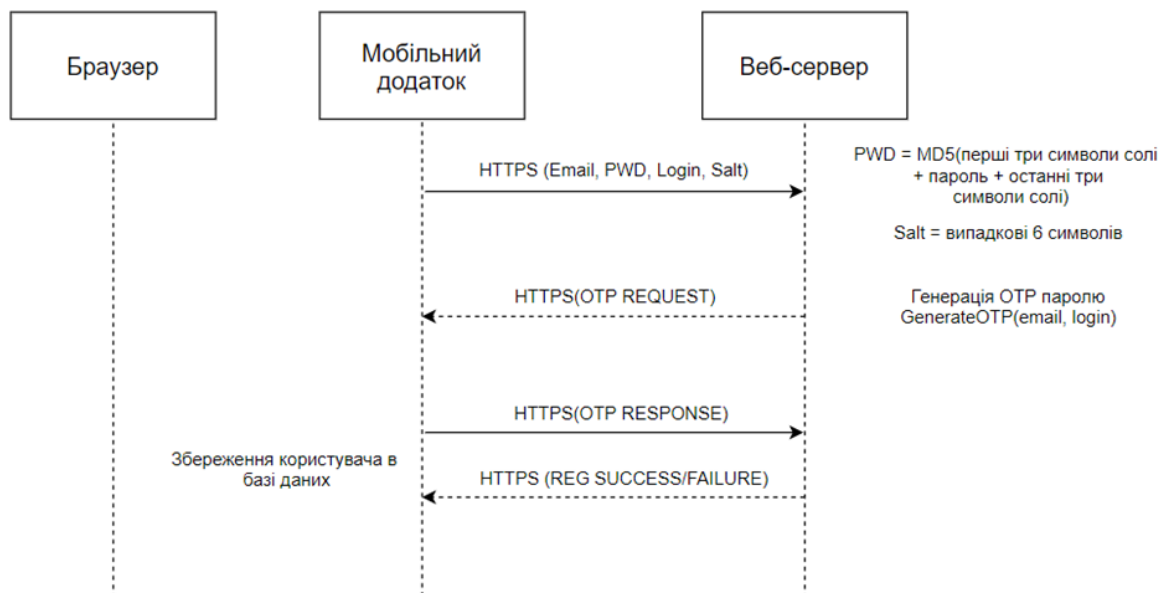


Рис.2.1. Обмін повідомленнями при реєстрації

3. Мобільний додаток за допомогою захищеного протоколу HTTPS відправляє дані реєстрації на сервер.

4. Сервер генерує тимчасовий пароль (OTP) та відправляє його на поштову адресу, вказану при реєстрації.

5. Користувач вводить на екрані смартфона отриманий пароль, після чого додаток відправляє дані на сервер.

6. Сервер надсилає відповідь про успішну, чи не успішну реєстрацію.

## 2.4 Використання логіну користувача

Пропонований протокол автентифікації використовує три об'єкти для входу користувача на сайт. Використовувані об'єкти: мобільний додаток, браузер і веб-сервер. Користувачеві також знадобиться веб-камера для входу в систему. Щоб увійти в обліковий запис веб-сайту, користувач повинен увійти в мобільний додаток веб-сайту. Причина в тому, що повідомлення відправляється в додаток, в якому користувач увійшов в систему. Для входу в мобільний додаток користувач вводить ідентифікатор користувача (login) і пароль (pwd), які відправляються на веб-сервер для підтвердження. Оскільки більшість користувачів не виходять з свого облікового запису, шанси на повторний вхід в мобільний додаток менше.

Процедура входу наступна:

1. Користувач спочатку відкриває веб-сторінку входу в систему в браузері.
2. На сторінці входу на веб-сайт відображається QR-код, що містить токен сеансу, створений веб-сервером.
3. Користувач сканує QR-код, що відображається на веб-сайті, за допомогою камери свого мобільного телефону.
4. Після отримання QR-коду мобільним додатком токен сеансу витягується з QR-коду.
5. Користувачу відображається на екрані повідомлення з IP-адресою користувача, що почав процес автентифікації.

6. Після підтвердження автентифікації на сервер відправляються дані, що містять: ID веб-сесії, ID-мобільної сесії, imei адреса телефону.
7. В браузері з'являється спливаюче вікно із запитом дозволу на доступ до веб-камери клієнтського комп'ютера. Користувач повинен дозволити доступ, щоб увійти на сайт.
8. Потім веб-камера робить знімок користувача та відправляє це зображення на сервер.
9. Додаток постійно оновлює статус мобільної сесії в очікуванні надходження зображень з веб-камери.
10. Після отримання зображення додатком його буде відображено на екрані та з'явиться пропозиція підтвердити або скасувати спробу автентифікації.
11. На сервер відправляється відповідь з параметрами сесії.
12. Сторінка веб-сайту автоматично оновлюється після успішної автентифікації.

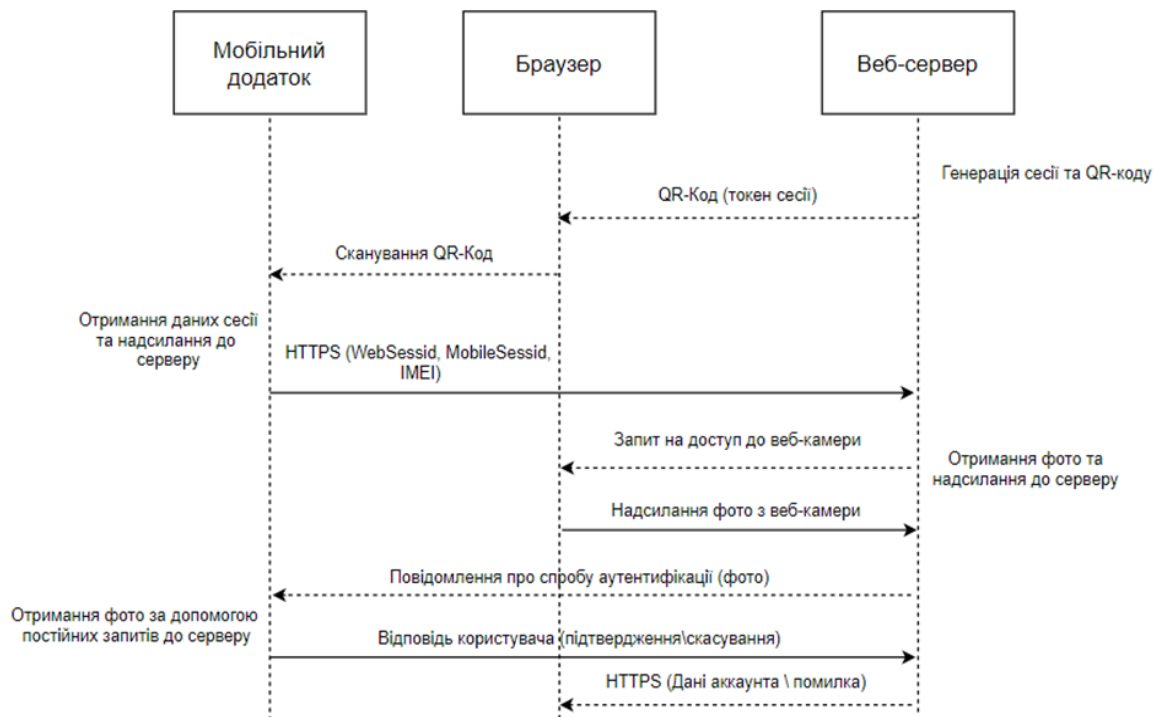


Рис.2.2. Обмін повідомленнями при автентифікації

Сеанс входу в систему існує протягом певного періоду часу, після чого термін дії сеансу закінчується, і веб-сервер перенаправляє користувача назад на домашню сторінку, відображаючи новий QR-код і, таким чином, не дозволяючи зловмисникові використовувати зображення, зроблене веб-камерою, в більш пізній етап автентифікації (рис. 2.2).

### 3 Тестування та оцінка працездатності схеми автентифікації

#### 3.1 Опис випробувальної установки

1. Смартфон OnePlus 5T з набором мікросхем Qualcomm Snapdragon 835 MSM8998, восьмиядерним процесором (чотири ядра Kryo 280 Performance з тактовою частотою 2,35 ГГц і чотири ядра Kryo 280 Efficiency з тактовою частотою 1,90 ГГц), 6 ГБ оперативної пам'яті і операційна система Android Pie з OxygenOS версії 9.0.4.

2. Настільний комп'ютер, який використовується для етапу входу в систему через Інтернет, Клієнтський комп'ютер є ПК з 64-розрядною операційною системою Windows 10 Professional з Google Chrome (версія 228.14.88), веб-камерою і процесором Intel® Core™ i7-7700K @ 4 ГГц з 64 ГБ оперативної пам'яті.

3. Веб-сайт для тестування етапу реєстрації і входу в систему розміщувався на настільному комп'ютері під керуванням 64-розрядної операційної системи Windows 10 на процесорі Intel® Core™ i7-3770 з тактовою частотою 3,40 ГГц і 8 Гб оперативної пам'яті.

4. Випробувальний веб-сайт було написано на ASP.NET Core і розміщено на сервері Windows IIS.

5. Сервер Microsoft SQL версії 17 використовувався для зберігання інформації користувача на стороні сервера.

### 3.2 Результати тестування

Для тестування на мобільному пристрої був запущений розроблений додаток. На головному екрані є вибір з 4-х кнопок (рис. 3.1): Логін (Login), Реєстрація (Register), Сканування QR (Scan QR), Деавторизація (Logout). Враховуючи перший запуск, на екрані доступні тільки кнопки логіну та реєстрації.

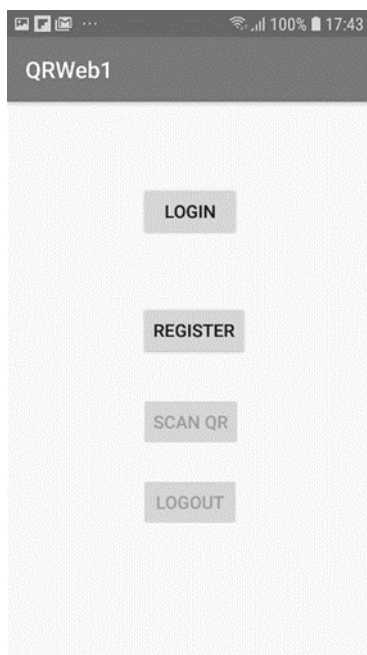


Рис.3.1. Основний екран додатку

Оскільки для тестування було вибрано «чистий» сервер без записів у БД, тож було обрано режим реєстрації та введено e-mail, логін та пароль (рис. 3.2).

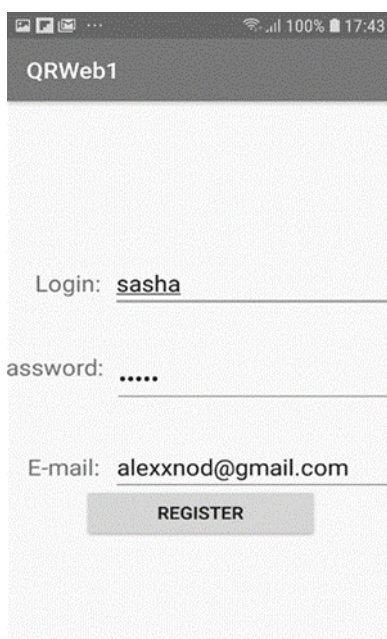


Рис.3.2. Екран реєстрації

Після натискання кнопки «register» на сервер було відправлено дані реєстрації. Після чого з'явився екран вводу тимчасового паролю (рис. 3.4), який був відправлений на адресу електронної пошти (рис. 3.3).



Рисунок 3.3. E-mail повідомлення з OTP кодом

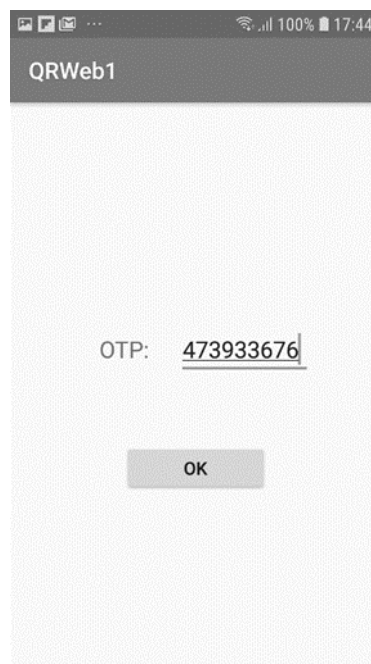


Рисунок 3.4. Вікно вводу OTP коду

Після вводу OTP коду було отримано повідомлення про успішну активацію (рис. 3.5).

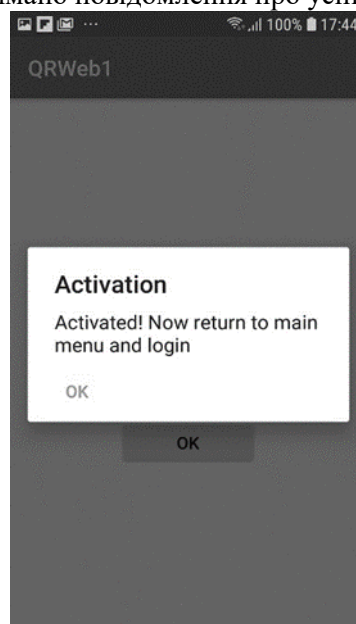


Рисунок 3.5. Повідомлення про успішну активацію

Після активації додатком було запропоновано перейти до вікна авторизації, де було введено логін та пароль (рис. 3.6).

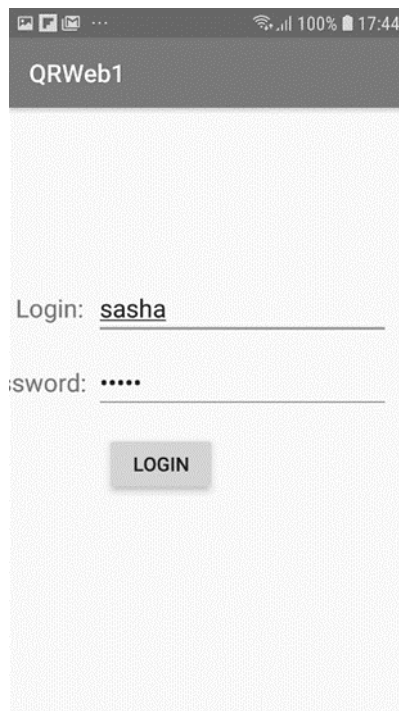


Рисунок 3.6. Вікно вводу логіну та паролю

Далі після успішної перевірки логіну та паролю додаток відкрив вікно вводу OTP паролю (див. рис. 2.2) який було направлено на e-mail, що був вказаний при реєстрації. Після проходження перевірки дані сесії були збережені у захищене сховище Android, а на екрані було відображено повідомлення про успішну авторизацію та пропозицію до переходу на основний екран і сканування QR-коду (рис. 3.7).

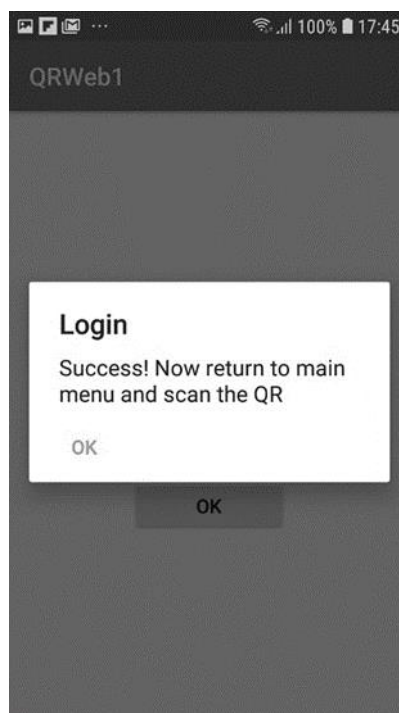


Рисунок 3.7. Повідомлення про успішну авторизацію



Далі було відкрито веб-браузер з необхідним тестовим сайтом (у локальній мережі по протоколу HTTPS). На екрані було відображено QR-код та текст, у якому пропонується відсканувати код за допомогою мобільного додатку (рис. 3.8).

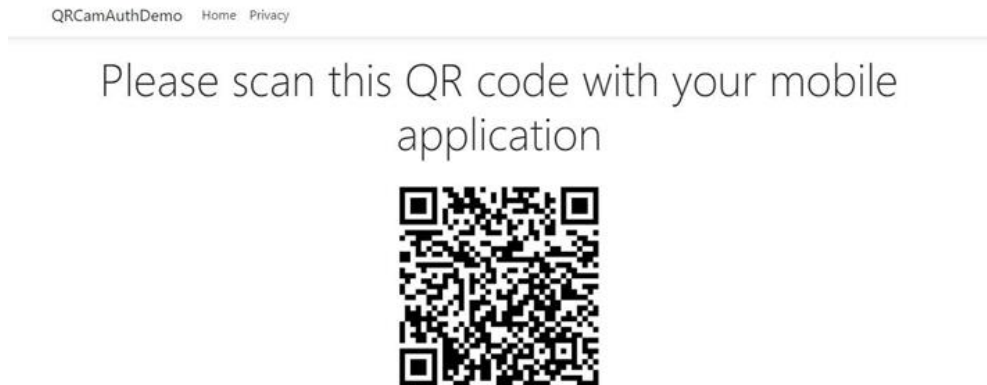


Рисунок 3.8. Веб-сторінка з QR-кодом

За допомогою додатка було відскановано даний QR-код, після чого на екрані було відображено інформацію про авторизацію, що включає в собі IP-адресу (рис. 3.9).

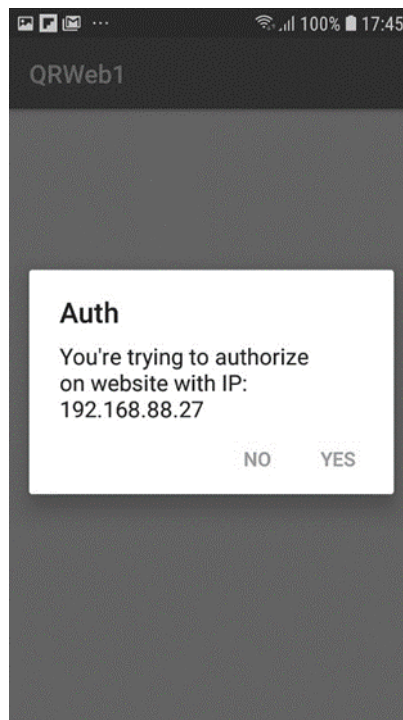


Рисунок 3.9. Запит на підтвердження авторизації

Після підтвердження авторизації сторінка з QR-кодом автоматично оновлюється і запитується доступ до веб-камери. Після надання дозволу на використання веб-камери пропонується зробити фотознімок користувача (рис. 3.10).

Please make the photo for further verification

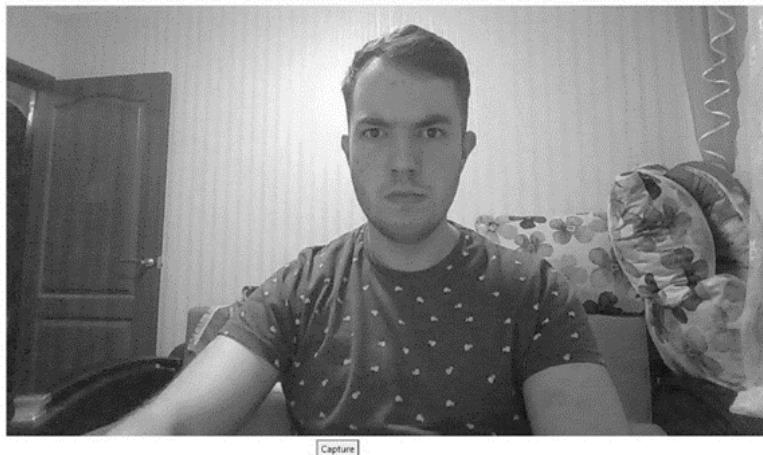


Рисунок 3.10. Вікно зображення веб-камери

Далі а на веб-сайті було відображено повідомлення про очікування підтвердження (рис. 3.11), а на телефон було відправлено повідомлення з відповідним зображенням (рис. 3.12)

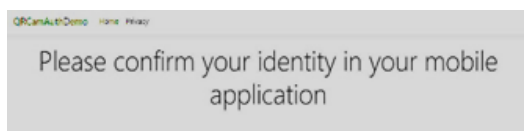


Рисунок 3.11. Повідомлення очікування підтвердження

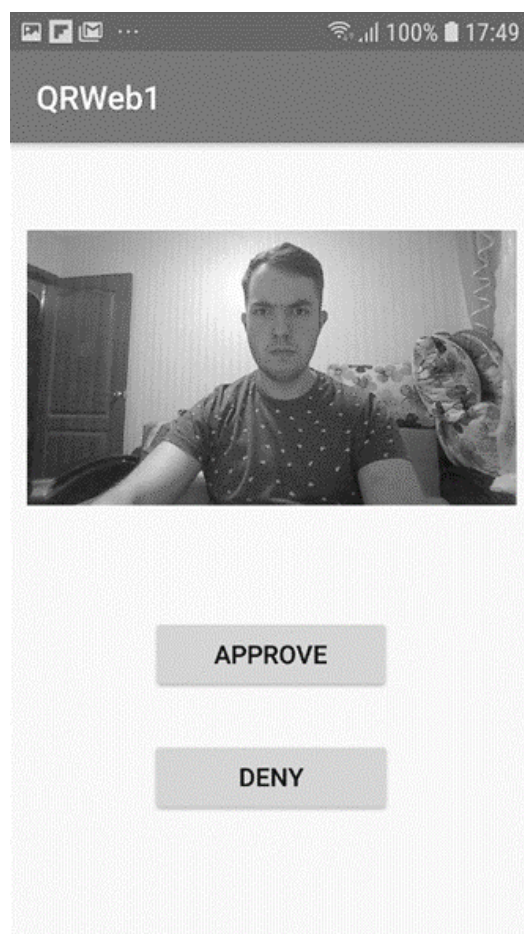


Рисунок 3.12. Вікно підтвердження у мобільному додатку



Після підтвердження успішної автентифікації у мобільному додатку сторінка веб-сайту оновлюється і відображається інформація про це (рис. 3.13).

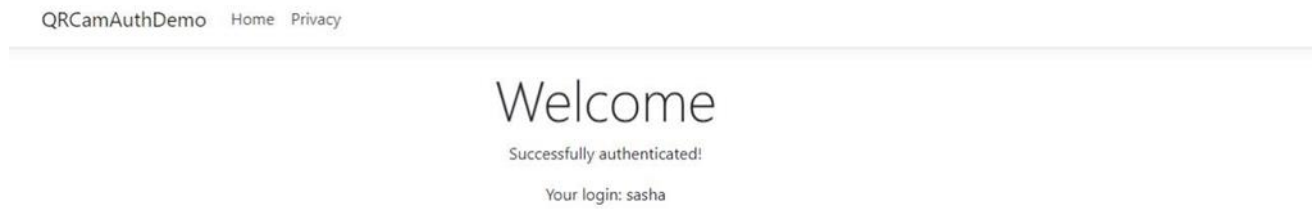


Рисунок 3.13. Повідомлення на веб-сайті про успішну авторизацію  
На цьому процедура автентифікації користувача завершується.

### 3.3 Розрахунок часу реєстрації та аналіз необхідних ресурсів

Час, необхідний для реєстрації користувачів ( $T_{REG}$ ), може бути розраховано на основі наступного виразу:

$$T_{REG} = T_{OTP} + T_{EC} + T_{MD5} + T_{SALT} + T_{DB} \quad (3.1)$$

де  $T_{OTP}$  – час генерування та доставки на електронну адресу тимчасового паролю,  $T_{EC}$  – час, витрачений користувачем на перевірку поштової скриньки на вміст коду,  $T_{MD5}$  – час генерування хешу паролю,  $T_{SALT}$  – час генерування солі,  $T_{DB}$  – час зберігання даних у захищеному сховищі Android.

Середній час витрачений на реєстрацію становить приблизно 38.25 секунд, якщо телефон під'єднано до мережі Wi-Fi, та 40.12 секунд, якщо телефон під'єднано до мережі LTE. Найбільшу частку часу займає перевірка e-mail.

Схожим чином розраховується час, необхідний користувачу для авторизації ( $T_{LOGIN}$ ) на веб-сайті при використанні запропонованого методу :

$$T_{LOGIN} = T_{FQR} + T_{SQR} + T_{SC} + T_{CAM} + T_{PN} + T_F \quad (3.2)$$

де  $T_{FQR}$  – час, необхідний для обробки та відображення QR-коду на веб-сайті,  $T_{SQR}$  – час, необхідний для сканування QR-коду через мобільний додаток,  $T_{SC}$  – час, необхідний для перевірки сесії мобільного додатку,  $T_{CAM}$  – час, необхідний для виконання знімку з веб-камери,  $T_{PN}$  – час, необхідний для надсилання та підтвердження фото у мобільному додатку,  $T_F$  – час, необхідний на проведення автентифікації та авторизації після підтвердження фото.

В середньому користувачеві потрібно 14,744 с для входу на веб-сайт за запропованою схемою, коли смартфон підключений через Wi-Fi, тоді як для підключення смартфона до мережі 4G потрібно 15,222 с.

Ресурси, що використовуються Android-додатком за запропованою схемою, визначаються за допомогою інструменту профілювання Android Studio. Зареєстровані значення мінімальної і максимальної завантаження ЦП на етапах реєстрації та входу в систему складають 1,2-19,8% і 1,4-27,3% відповідно. Середній обсяг пам'яті, який використовується Android-додатком для запропонованої схеми, становить 46 МБ. На основі аналізу одержаної статистики можна зробити висновок, що вимоги до використання ЦП і пам'яті для запропонованої схеми досить невеликі, і тому запроповану схему можна використовувати для входу на веб-сайти без істотних обмежень.

### 3.4 Порівняння з відомими схемами на основі критерію зручності використання

Запропонована схема була порівняна з деякими існуючими схемами по критерію зручності використання. Зручність використання схеми автентифікації враховує кількість токенів, необхідних схемою автентифікації, і кількість токенів автентифікації, які користувач повинен запам'ятати. Це порівняння також бере до уваги будь-яке додаткове програмне або апаратне

забезпечення. Смартфон потрібний для більшості існуючих схем, оскільки його мають всі користувачі Інтернету.

Деякі з існуючих схем вимагають, щоб на клієнтському комп'ютері був встановлений конкретний модуль або драйвер, апаратний токен або довірена третя сторона і т. д. Пропонована схема вимагає тільки наявності комп'ютерної камери, яка легко доступна на всіх настільних комп'ютерах, особливо портативних. Крім того, пропонована схема не вимагає, щоб користувач запам'ятав будь-який токен аутентифікації, і, таким чином, вона більш зручна для користувача в порівнянні з іншими схемами аутентифікації.

Порівняння, засноване на критерії зручності використання, також розглядає потребу в доступі до Інтернету на телефоні як одному з факторів для вимірювання зручності використання в різних схемах автентифікації для розуміння витрат, понесених при використанні схеми. Пропонована схема вимагає наявності доступу до Інтернету на телефоні, який став синонімом смартфонів. Таблиця 2 описує порівняльний аналіз, заснований на зручності використання.

### 3.5 Порівняння на основі оцінки безпеки застосування

В цьому розділі запропонована схема порівнюється з відомими існуючими схемами на основі критерію безпеки застосування, яку вони забезпечують проти атак, описаних в розділі 2.2.

#### 1. Фішингові атаки RT MITM і CR MITM:

1.1. U-PWD [23, 24] і схеми аутентифікації на основі OTP / PIN, такі як Google 2-step [7] і SAASPASS [8], уразливі як для RT MITM, так і для CR MITM фішингових атак.

Таблиця 3.1. Порівняння за кількістю використовуваних токенів і їх безпеки

№	Схема	Токени, що використовуються схемою	Токени, що необхідно запам'ятати	Додаткові потреби	Необхідність в інтернеті на телефоні
1	Google 2 Step [7]	3-U, PWD, OTP on SP	2-U, PWD	Мобільний телефон	Ні
2	SAASPASS [8]	3-U, PWD, OTP on App	2-U, PWD	Смартфон	Так
3	Xie et al. [16]	4-U, PWD, DH Public (g, p), Private Up	2-U, PWD	Веб-камера, смартфон	Так
4	Kim et al. [10]	4-U, PWD, Session ID, Secret Key	2-U, PWD	Смартфон з GPS	Так
5	Mukhopadhyay et al. [11]	3-U, PWD, Secret Key in SP	2-U, PWD	Смартфон	Так
6	Dodson et al. [12]	4-U, PWD, Secret Key, QR-Code	0-NIL (Сканування QR-коду)	Смартфон	Так
7	Leung et al. [13]	4-U, PWD, Secret Key, OTP CAPTCHA	2-U, PWD	Відсутні	Невідомо
8	Zhu et al. [22]	3U, SALT, PWD, CAPTCHA	2- U, PWD	Відсутні	Невідомо
9	Tricipher [19]	3U, PWD, TPM Secret Key, TACS credential	2- U, PWD	SAPI драйвер, окреме обладнання	Невідомо
10	RSA SecurID HW Token [21]	4-U, PWD, HW token information, PIN	2-U, PWD	Окреме обладнання	Невідомо
11	Yubikey U2F [14]	5-KPUB, KPRIV, Counter, U, PWD	2-U, PWD	Окреме обладнання	Невідомо

12	Push Login [15, 20]	3-U, PWD, SP	1-U	Смартфон	Так
13	Password Managers [17, 18]	3-U, PWD, the master key of the password manager	Master PWD	Відсутні	Невідомо
14	U-PWD [23, 24]	2-U, PWD	2-U, PWD	Відсутні	Невідомо
15	Запропонована схема	3-U, фото зроблене за допомогою веб-камери	0-NIL (Сканування QR-коду)	Смартфон, веб-камера	Так

- 1.1. Зловмисник може легко отримати реєстраційні дані користувача за допомогою фішингового сайту. Зловмисник може обманути користувача, показавши йому точну копію справжнього веб-сайту, і передати облікові дані, введені користувачем на фішинговому веб-сайті, в режимі реального часу на справжній веб-сайт. Зловмисник також може запустити фішингову атаку CR MITM за схемою на основі OTP / PIN, просто перенаправив свій віддалений робочий стіл через клієнтський термінал, і користувач буде спокушений ввести облікові дані, які фактично вводяться на віддаленому робочому столі зловмисника.
- 1.2. Схеми на основі QR-коду також не можуть протистояти атакам RT MITM і CR MITM. Схема, запропонована Xie et al. [16] може бути атакована за допомогою підробленого шкідливого розширення браузера. Зловмисник може встановити на комп'ютер користувача підроблене шкідливе розширення браузера. Таким чином зловмисник отримає облікові дані, які зловмисник передасть розширенню SamAuth. Розширення SamAuth відправить інформацію про користувача на сервер, а також ініціює обмін даними по типу схеми Діффі-Хеллмана. Сервер перевіряє облікові дані, і таким чином обліковий запис користувача буде відправлено сервером в браузер зловмисника. Однак схема Се і ін. [16] захищена від атаки CR MITM, оскільки зловмисник не може отримати доступ до камери ПК користувача. Точно також схема на основі QR-коду, запропонована Kim et al. [10], не захищена від атак RT MITM і CR MITM. Причина в тому, що IP-адреса, присутня в QR-коді, може бути підроблена, і зловмисник може відправити запит до серверу. Схема, запропонована Mukhopadhyay et al. [11] також вразлива для подібних атак, оскільки облікові дані, введені користувачем, можуть бути легко отримані зловмисником за допомогою фішингового веб-сайту, а QR-код може бути переданий на фішинговий веб-сайт. Таким чином, процес входу в систему завершується, і зловмисник отримує доступ до облікового запису користувача. Точно також схема Додсона і ін. [12] є вразливою для RT MITM, а також для атаки CR MITM, оскільки зловмисник передає QR-код на фішинговий веб-сайт, який сканується смартфоном користувача.
- 1.3. Схеми на основі апаратних токенів, такі як Tricipher [19] і Yubikey з використанням U2F [20], захищені від фішингових атак RT MITM і CR MITM через використання складових облікових даних. Однак програмний / апаратний токен RSA SecurID [21] не забезпечує від цих атак, оскільки зловмисник може отримати код доступу RSA через фішинговий веб-сайт.
- 1.4. Графічна схема автентифікації на основі пароля, запропонована Leung et al. [13] захищена від атаки RT MITM, тому що ретрансляція координат клацання мишею користувача при переміщенні CAPTCHA досить складна, і може бути значущою різниця в розширенні екрану робочого столу зловмисника і робочого столу користувача. Однак ця схема не забезпечена від атаки CR MITM. Схема Чжу і ін. [14] не забезпечена від RT MITM, а також від CR MITM-атак, тому що зловмисник може записувати і відображати координати клацання користувача на CaRP, який відображається на справжньому веб-сайті.

- 1.5. Схеми входу в систему на основі push-повідомлень [15, 16] уразливі для атак RT MITM і CR MITM, оскільки зловмисник може ретранслювати введені облікові дані користувачем, і потім користувач затвердить push-повідомлення, отримане в додатку для смартфона.
- 1.6. Менеджери паролів [17, 18] захищені від атак RT MITM і CR MITM, оскільки облікові дані користувача автоматично відправляються менеджерами паролів, і користувачеві не потрібно їх вводити. Однак, якщо зловмисник вводить неправильні облікові дані, він запропонує користувачеві повторно ввести своє ім'я і пароль, що призведе до можливості реалізації атаки RT MITM і CR MITM.
- 1.7. Пропонована схема захищає користувача від фішинг-атаки RT MITM, оскільки навіть після ретрансляції QR-коду зловмисник не зможе відправити на сервер фотографію користувача, зроблену його веб-камерою (зрозуміло, що зловмисник не має фізичного доступу до пристрою користувача). Причина в тому, що фотографія повинна бути зроблена веб-камерою зловмисника. Більш того, після того, як користувач сканує QR-код з програми для смартфона, веб-камера зловмисника відразу ж зробить фото зловмисника за допомогою його веб-камери і відправить цю фотографію користувачеві через push-повідомлення, тим самим попередивши користувача про спробу підміни. Пропонована схема також захищена від фішингових атак CR MITM, оскільки зловмисник не може отримати доступ до веб-камери користувача.

## 2. Фішингові атаки на основі шкідливих розширень браузера.

Фішингові атаки на основі шкідливих розширень браузера включають ведення кейлогерів, ведення журналу екрану і перехоплення пароля. Зловмисник може зламати двоетапний токен Google [13], U-PWD [24], SAAS-PASS [8], програмний / апаратний [21] токен RSA SecurID, використовуючи кейлогінг і сніфінг паролів, оскільки у вищезгаданих схемах всі облікові дані вводяться на веб-сайті, і, таким чином, зловмисник може перехопити їх через шкідливе розширення браузера. Зловмисник також може зламати менеджери паролів, встановивши на клієнтському терміналі шкідливе розширення браузера, яке може перехопити пароль перед його відправкою на сервер. Схема Xie et al. [9] захищена від кейлоггерів і можливості перехоплення паролів, оскільки шкідливе розширення браузера не може отримати доступ до інформації, введеної користувачем в розширенні CamAuth через політику одного і того ж походження. Однак, зловмисник може отримати доступ до QR-коду за допомогою запису екрану, але не зможе порушити повну автентифікацію. Kim et al. [10] і Dodson et al. [12] захищають користувачів від цих атак, оскільки не вимагають введення будь-яких облікових даних. Схема Leung і ін. на основі CAPTCHA [13] захищені від кейлоггерів і перехоплення паролів, оскільки використовується OTP CAPTCHA на основі флеш-пам'яті. З іншого боку, зловмисник може зламати схему Чжу і ін. [14] і отримати введення користувача через ведення журналу екрану. Оскільки Zhu et al. [14], Tricipher [19], Mukhopadhyay et al. [11] і схеми на основі примусового входу в систему [15, 16] використовують або довірена пристрій, або другий фактор автентифікації, зловмисник може отримати тільки токен автентифікації або ідентифікацію користувача, але не може порушити функціонування цієї схеми. Пропонована схема захищена від подібних атак, оскільки користувачеві не потрібно вводити свої облікові дані, і тому шкідливі розширення браузера не зможуть перехопити цю інформацію.

Таблиця 3.2. Порівняння з точки зору захисту від відомих загроз

№	Схема	RT MITM	CR MITM	Кейлогінг	Запис екрану	Сніфінг паролю	Підміна додатків	Рівень безпеки
1	Google 2 step	-	-	-	-	-	-	0
2	SAASPASS	-	-	-	-	-	-	0

3	Xie et al.	-	+	+	-	+	-	3
4	Kim et al.	-	-	+	+	+	-	3
5	Mukhopadhyay et al.	-	-	-	-	-	+	1
6	Dodson et al.	-	-	+	+	+	+	4
7	Leung et al.	+	-	+	-	+	-	3
8	Zhu et al.	-	-	-	-	-	+	1
9	Tricipher	+	+	-	-	-	+	3
10	RSA SecurID I-IW token	-	-	-	-	-	-	0
11	Yubikey U2F	+	+	+	-	-	+	4
12	Push login	-	-	-	-	-	-	0
13	Password Managers	+	+	+	-	+	+	5
14	U-PWD	-	-	-	-	-	-	0
15	Запропонована схема	+	+	+	+	+	+	6

**3. Підміна додатків:** зловмисник може встановити підроблений мобільний додаток або розширення на смартфон, або робочий стіл користувача, щоб отримати інформацію про нього. Xie et al. [9], Google 2-step [7], SAASPASS [8], U-PWD, Kim et al. [10] уразливі для цієї атаки, оскільки підроблений додаток може бути встановлено на пристрій користувача, а облікові дані, введені ним, зберігаються в підробленому додатку, який встановлено зловмисником. Zhu et al. Схема [14] захищена від цієї атаки, оскільки не використовує ніяких додатків або розширень. Mukhopadhyay et al. [11], Tricipher [19] і Dodson et al. [12] захищені від подібної атаки, оскільки в цих схемах принаймні одна частина токена аутентифікації зберігається на пристрої користувача і не вводиться в додаток. Пропонована схема також захищена від спуфінга додатків, оскільки секретні дані зберігаються в додатку, зашифрованому за допомогою Android Keystore API, який не буде доступний в підробленому додатку. Порівняння запропонованої схеми з іншими існуючими схемами, заснованими на безпеці застосування представлено в таблиці 3.

### 3.5 Висновки

В роботі було запропоновано нову схему автентифікації, яка здатна протистояти як традиційним, так і ускладненим фішинговим атакам. Пропонований протокол автентифікації не тільки здатний впоратися з атаками типу RT MITM і CR MITM, але також може захистити користувачів від дії шкідливих розширень браузера і підміни автентичних додатків на шкідливі їх варіанти. Окрім цього, запропонована схема не вимагає від користувачів необхідності мати будь-які токени автентифікації або облікові дані, оскільки їм просто потрібно відсканувати QR-код і перевірити зображення, зняте власною веб-камерою. Це робить використання запропонованої схеми більш зручною і простою для користувачів в порівнянні з іншими відомими схемами автентифікації.

Наразі застосування запропонованої схеми вимагає використання веб-сайтами протоколу HTTPS для обміну всіма задіяними в обміні даними. Таким чином, запропонований протокол може бути реалізований для безпечного управління файлами cookie, щоб він не був вразливий для перехоплення даних сеансу. У майбутньому доцільно провести більш широке дослідження, направлене на оцінку зручності використання користувачем запропонованої схеми. Також, можливо, представлять інтерес визначення в майбутньому таких параметрів її використання, як

крива навчання і масштабованість, а також оцінки дослідження можливостей запропонованої схеми в обробці великої кількості одночасних запитів.

#### ЛІТЕРАТУРА

1. Ластдрагер Е.Е. Досягнення консенсусного визначення фішингу на основі систематичного огляду літератури. *Crime Sci.* 2014. Вип. 3. С.21-32.
2. Абдельхамід Н. Багатознакові правила класифікації фішингу. *Прикладні обчислення та інформатика.* 2015. Вип. 11. С. 29-46.
3. Бандей М.Т., Кадрі Ж.А. Фішинг - зростаюча загроза для електронної комерції, *The Business Review.* 2011. Вип. 12. С. 76-83.
4. Бадра М., Ель-Сауда С., Хадже І., *Фішинг-атаки та рішення: матеріали 3-ї Міжнародної конф. з мобільних мультимедійних комунікацій, MobiMedia 2007, Нафпактос, Греція, 27-29 серпня 2007 р.* С. 42-43.
5. Рамі М. М., Фаді Т., Лі М. Підручник та критичний аналіз методів фішингових веб-сайтів, *Computer Science Review.* 2015. Вип. 17. С. 1-24.
6. Джагатич Т., Натаніель Дж., Менцер Ф. Соціальний фішинг, *Commun.* 2007. С. 94–100.
7. Посилений захист вашого облікового запису Google. Google: веб-сайт. URL: <https://www.google.com/landing/2step/> (дата звернення: 19.10.2020)
8. Багатофакторна автентифікація. SAASPASS: веб-сайт. URL: <https://saaspass.com/> (дата звернення: 20.10.2020)
9. Захищена схема автентифікації, щоб зірвати RT MITM, CR MITM та фішинг-атаки на основі зловмисного браузера. Гаурав В., Маной М., Прадіп А.: веб-сайт. URL: <https://www.sciencedirect.com/science/article/abs/pii/S2214212618300140> (дата звернення: 20.10.2020)
10. Син-Хюн К., Десон К., Син-Хун Дж., Сунг-Хун Л., *Схема автентифікації на основі геолокації на основі QR-коду, щоб перемогти активну фішинг-атаку в режимі реального часу: матеріали семінару з управління цифровими ідентичностями (DIM '13). Асоціація обчислювальних машин, Нью-Йорк, Нью-Йорк, США, 2013 р.* С. 51–62.
11. Мухопадхьяй С., Арглес Д., *Механізм боротьби з фішингом для єдиного входу на основі QR-коду: матеріали міжнародної конференції з інформаційного суспільства (i-Society 2011), Лондон, 2011 р.* С. 505-508.
12. Додсон Б., Сенгупта Д., Боне Д., Лам М.С., *Безпечна, зручна для споживачів веб-автентифікація та платежі за допомогою телефону: матеріали із збір. MobiCASE 2010, конспекту лекцій Інституту комп'ютерних наук, Спрінгер, Берлін, Гейдельберг, 2010. Т. 76.*
13. Лун Ч., *Придушити фішинг за допомогою CAPTCHA за допомогою OTP: матеріали третьої міжнародної конференції з питань боротьби з піддробкою, безпекою та ідентифікацією у зв'язку, Гонконг, 2009.* С. 187-192.
14. Ваш ключ до найшвидшого та найбезпечнішого входу. Yubico: веб-сайт. URL: <https://www.yubico.com/why-yubico/for-individuals/> (дата звернення: 21.10.2020)
15. Г. Варшні та М. Місра, *Вхід на основі push-повідомлень із використанням пристроїв BLE: матеріали 2-ї міжнародної конференції з інформаційних технологій, інформаційних систем та електротехніки (ICITISEE), Джокьякарта, 2017.* С. 479-484.
16. М. Се, Й. Лі, К. Йошиго, Р. Секер и Дж. Біан, *Забезпечення автентифікації в Інтернеті за допомогою камери: матеріали 16-го Міжнародного симпозиуму з інженерії систем високих гарантій, Дейтона-Біч-Шорз, Флорида, 2015.* С. 232-239.
17. Lastpass запам'ятовує всі ваші паролі. Lastpass: веб-сайт. URL: <https://www.lastpass.com/> (дата звернення: 21.10.2020)
18. Росс Б., Джексон С., Міяке Н., Боне Д., Мітчелл Дж. *Потужніша автентифікація пароля за допомогою розширення браузера.* Балтімор, США, 2005. С. 17-32.
19. Довідковий документ про запобігання фішинг-атакам людини із багатофакторною аутентифікацією. Tricipher: веб-сайт. URL: <https://www.globaltrustit/documents/press/phishing/PhishingSolutionWhitepaper.pdf> (дата звернення: 22.10.2020)

20. Увійти в систему Yahoo. Yahoo: веб-сайт. URL: <https://login.yahoo.com/> (дата звернення: 23.10.2020)
21. RSA SecurID. Veal V: веб-сайт. URL: [https://www.webopedia.com/TERM/R/rsa\\_secure\\_id.html](https://www.webopedia.com/TERM/R/rsa_secure_id.html) / (дата звернення: 23.10.2020)
22. Б.Б. Чжу, Дж. Ян, Г. Бао, М. Ян і Н. Сю, *Капча як графічні паролі - новий примітив безпеки, заснований на складних проблемах III*, у "Протоколах угод з інформаційної криміналістики та Безпеки, т. 9, №6, червень 2014 р. С. 891-904.
23. Дж. Бонно і С. Прейбуш, *Зарості паролів: технічні та ринкові помилки при аутентифікації людини в Інтернеті*, в Proc. WEIS. 2010 р. С. 1-48.
24. Дж. Бонно, К. Герлі, П.К. Ооршота та Ф. Стаджано, *Прагнення замінити паролі: рамки для порівняльної оцінки схем автентифікації через Інтернет: матеріали симпозиуму з безпеки та конфіденційності*, Сан-Франциско, СА, 2012 р. С. 553-567.

## REFERENCES

1. E.E. Lastdrager. Achieving a consensual definition of phishing based on a systematic review of the literature, *Crime Sci* vol. 3, 2014, pp. 21-32, DOI:<https://doi.org/10.1186/s40163-014-0009-y> [in English]
2. A. Neda. Multi-label rules for phishing classification, *Applied Computing and Informatics*, vol. 11, 2015, pp. 29-46, DOI:<https://doi.org/10.1016/j.aci.2014.07.002> [in English]
3. Banday M.T., Qadri J.A. Phishing - A Growing Threat to E-Commerce, *The Business Review*, vol. 12, 2011, pp. 76-83. [in English]
4. Badra M., El-Sawda S., Hajjeh I., "Phishing attacks and solutions", in *Proceedings of the 3rd International Conference on Mobile Multimedia Communications, MobiMedia 2007*, Nafpaktos, Greece, August 27-29, 2007, pp. 42-43. [in English]
5. Rami M. M., Fadi T., Lee M., Tutorial and critical analysis of phishing websites methods, *Computer Science Review*, vol. 17, 2015, pp. 1-24, DOI:<https://doi.org/10.1016/j.cosrev.2015.04.001> [in English]
6. Jagatic T., Nathaniel J., Menczer F. Social phishing, *Commun*, 2007, pp. 94-100. [in English]
7. Stronger security for your Google Account. Google., 2015. [Online]. Available: <https://www.google.com/landing/2step/>. [Accessed October 19, 2020]. [in English]
8. Multi-Factor Authentication. SAASPASS., 2019. [Online]. Available: <https://saaspass.com/>. [Accessed October 20, 2020]. [in English]
9. Secure authentication scheme to thwart RT MITM, CR MITM and malicious browser extension based phishing attacks. Gaurav V., Manoj M., Pradeep A., 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S2214212618300140>. [Accessed October 20, 2020]. [in English]
10. Seung-Hyun K., Daeseon C., Seung-Hun J., Sung-Hoon L., "Geo-location based QR-Code authentication scheme to defeat active real-time phishing attack", in *Proceedings of the 2013 ACM workshop on Digital identity management (DIM '13)*. Association for Computing Machinery, New York, NY, USA, 2013, pp. 51-62. DOI:<https://doi.org/10.1145/2517881.2517889> [in English]
11. Mukhopadhyay S., Argles D., "An Anti-Phishing mechanism for single sign-on based on QR-code", in *Proceedings of the International Conference on Information Society (i-Society 2011)*, London, 2011, pp. 505-508, DOI:10.1109/i-Society18435.2011.5978554 [in English]
12. Dodson B., Sengupta D., Boneh D., Lam M.S., "Secure, Consumer-Friendly Web Authentication and Payments with a Phone", in *Proceedings of the MobiCASE 2010. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol 76. Springer, Berlin, Heidelberg. DOI:[https://doi.org/10.1007/978-3-642-29336-8\\_2](https://doi.org/10.1007/978-3-642-29336-8_2) [in English]
13. Leung C., "Depress phishing by CAPTCHA with OTP", in *Proceedings of the 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication*, Hong Kong, 2009, pp. 187-192, DOI:10.1109/ICASID.2009.5276926. [in English]



14. Your key to the fastest, safest login. Yubico., 2019. [Online]. Available: <https://www.yubico.com/why-yubico/for-individuals/>. [Accessed October 21, 2020]. [in English]
15. G. Varshney and M. Misra, "Push notification based login using BLE devices", in *Proceedings of the 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, Yogyakarta, 2017, pp. 479-484, DOI: 10.1109/ICITISEE.2017.8285554. [in English]
16. M. Xie, Y. Li, K. Yoshigoe, R. Seker and J. Bian, "CamAuth: Securing Web Authentication with Camera", in *Proceedings of the 16th International Symposium on High Assurance Systems Engineering*, Daytona Beach Shores, FL, 2015, pp. 232-239, DOI:10.1109/HASE.2015.41. [in English]
17. Lastpass remembers all your passwords. Lastpass., 2019. [Online]. Available: <https://www.lastpass.com/>. [Accessed October 21, 2020]. [in English]
18. Ross B., Jackson C., Miyake N., Boneh D., Mitchell JC., *Stronger password authentication using browser extension*. Baltimore, MD, USA, 2005, pp. 17-32. [in English]
19. White paper preventing man in the middle phishing attacks with multi-factor authentication. Tricipher., 2019. [Online]. Available: <https://www.globaltrustit/documents/press/phishing/PhishingSolutionWhitepaper.pdf>. [Accessed October 22, 2020]. [in English]
20. Yahoo sign in. Yahoo., 2016. [Online]. Available: <https://login.yahoo.com/>. [Accessed October 23, 2020]. [in English]
21. RSA SecurID. Beal V., 2019. [Online]. Available: [https://www.webopedia.com/TERM/R/rsa\\_secure\\_id.html/](https://www.webopedia.com/TERM/R/rsa_secure_id.html/). [Accessed October 23, 2020]. [in English]
22. B. B. Zhu, J. Yan, G. Bao, M. Yang and N. Xu, "Captcha as Graphical Passwords—A New Security Primitive Based on Hard AI Problems," in " ", in *Proceedings of the Transactions on Information Forensics and Security*, vol. 9, no. 6, pp. 891-904, June 2014, DOI:10.1109/TIFS.2014.2312547. [in English]
23. J. Bonneau and S. Preibusch, "The password thicket: technical and market failures in human authentication on the web," in *Proc. WEIS 2010*, pp 1-48. [in English]
24. J. Bonneau, C. Herley, P. C. v. Oorschot and F. Stajano, "The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes," in *Proceedings of the Symposium on Security and Privacy*, San Francisco, CA, 2012, pp. 553-567, DOI:10.1109/SP.2012.44. [in English]



UDK 517.977.5

## Mathematical model of condition-based preventive maintenance of a complex technical system

N.S. Podtcykin

**Podtcykin Nikolay**

*Associate professor*

*Karazin Kharkiv National University, Svobody Sq 4, Kharkiv, Ukraine, 61022*

*e-mail: [n.podtcykin@karazin.ua](mailto:n.podtcykin@karazin.ua)*

*<https://orcid.org/0000-0001-5814-9164>*

A new approach to mathematical modeling of complex technical systems according to their condition is being developed. Connections between subsystems of a complex system are considered to be arbitrary in terms of reliability. Due to wear, failures of subsystems can happen at random moments of time. Failures of some subsystems can lead to the entire system failure. The purpose of the simulation is to maintain the level of reliability and operability of a complex technical system at an optimal level for an unlimited time interval by means of regular preventive maintenance and repair. Technical instructions and specifications, as well as statistical data, are used in modeling a priori characteristics of subsystems. That information is used to determine the reliability of a complex system and its condition. The mathematical model is built in terms of the Markov decision-making process. The chosen optimization method allows obtaining the best policy for choosing acceptable preventive maintenance policy and repairs at the planned time of inspections and moments of failures.

*Key words: mathematical model, complex system, technical system wear, system operability, system state, control strategy.*

## Математична модель профілактики складної технічної системи за станом

М.С. Подцикін

**Подцикін Микола  
Серафимович**

*Кандидат технічних наук, доцент, Харківський національний університет імені В.Н. Каразіна, майдан Свободи 4, м. Харків, Україна, 61022*

*e-mail: [n.podtcykin@karazin.ua](mailto:n.podtcykin@karazin.ua)*

*<https://orcid.org/0000-0001-5814-9164>*

Розвивається новий підхід в математичному моделюванні складних технічних систем з урахуванням їх станів. Розглядаються складні системи, що складаються з кінцевого числа, взагалі кажучи, різних простих підсистем. Допускається, що в сенсі надійності підсистеми з'єднані між собою довільно. Усі підсистеми в процесі роботи зношуються, ймовірність їх відмови збільшується. Відмова однієї або декількох підсистем не обов'язково призводить до відмови усієї системи. Якщо відмова складної системи відбувається, то це призводить до великих втрат із-за її простою і витрат на відновлення. У статті розглянуто випадок роботи системи на необмеженому інтервалі часу з регулярними профілактичними обслуговуваннями і ремонтами в моменти відмов. Допускається безліч різних видів обслуговувань і ремонтів. Метою моделювання системи є знаходження оптимального правила вибору виду обслуговування в планові моменти контролю системи та виду ремонту в моменти відмов, з урахуванням спостережуваних станів, на необмеженому інтервалі часу. Технічні описи та статистичні дані визначають оцінку функції надійності кожної підсистеми. Функції надійності підсистем і інформація про їх з'єднання між собою в сенсі надійності визначають оцінку функції надійності складної системи. Стан системи в момент контролю визначається значеннями параметрів підсистем і функцією надійності системи. Таке визначення стану забезпечує оцінку ймовірності відмови і продуктивності в кожен момент контролю для конкретної розглянутої системи. Використання стану в моделі профілактики є актуальним, оскільки враховуються особливості, ступінь зносу конкретної системи. У існуючих моделях профілактики системи, побудованих на основі "напрацювання на відмову", для прийняття рішень з підтримки оптимального рівня надійності використовуються характеристики, отримані їх усередненням по цілому ансамблю однотипних систем. Зрозуміло, що при цьому особливості конкретної системи не враховуються. Вартість відновлення системи після відмови часто залежить від набору підсистем, які привели її до відмови. У пропонованій статті розглянуто можливість знаходження і обліку таких наборів. Модель побудована в термінах марківського процесу прийняття рішень. Пропонується використати відомий метод оптимізації, заснований на принципі стислих відображень. Цей метод забезпечує знаходження правила вибору профілактик і ремонтів для необхідного рівня

ефективності роботи системи. Деякі ключові моменти побудови моделі проілюстровані на конкретному прикладі складної системи.

**Ключові слова:** математична модель, складна система, знос технічної системи, працездатність системи, стан системи, стратегія управління.

## Математическая модель профилактики сложной технической системы по состоянию

Н.С. Подцыкин

**Подцыкин Николай Серафимович** *Кандидат технических наук, доцент, Харьковский национальный университет имени В.Н. Каразина, пл. Свободы 4, г. Харьков, Украина, 61022*  
*e-mail: [n.podtcikin@karazin.ua](mailto:n.podtcikin@karazin.ua)*  
*<https://orcid.org/0000-0001-5814-9164>*

Развивается новый подход в математическом моделировании сложных технических систем по состоянию. Подсистемы, рассматриваемой сложной системы, соединены между собой в смысле надежности произвольно. В результате износа допускаются отказы подсистем в случайные моменты времени. Отказы некоторых подсистем могут приводить к отказу всей системы. Целью моделирования является поддержание уровня надежности и работоспособности сложной технической системы на оптимальном уровне на неограниченном интервале времени с помощью регулярных профилактических обслуживаний и ремонтов. Технические инструкции и описания, а также статистические данные, использованы при моделировании априорных характеристик подсистем. Эта информация используется в модели при определении надежности сложной системы и ее состояния. Математическая модель построена в терминах марковского процесса принятия решений. Выбранный метод оптимизации позволяет получить наилучшую стратегию выбора допустимых профилактических обслуживаний и ремонтов в плановые моменты контроля и в моменты отказов. Некоторые ключевые и громоздкие элементы построения модели проиллюстрированы на примере выбранной сложной технической системы.

**Ключевые слова:** математическая модель, сложная система, износ технической системы, работоспособность системы, состояние системы, стратегия управления.

### 1 Introduction

The method for constructing a mathematical model of a complex technical system in order to optimize the level of its reliability and performance by means of preventive maintenance and repair for an unlimited time interval is discussed in the article.

The class of the systems under consideration is limited by the following conditions. The system consists of a finite number of subsystems. All subsystems in a complex system are used to achieve the same goal. All subsystems wear out during operation, their level of functional quality decreases, and the failure probability increases. Subsystem failure leads to its shutdown and following recovery costs. We assume that in terms of reliability the subsystems are interconnected arbitrarily and function independently of each other. Therefore, the failure of one or more of them does not necessarily lead to the complex system failure. But system failure leads to long downtime and could be very expensive. It follows that the substantiation of the optimal policy for choosing preventive maintenance at the planned inspection time and at the moments of system failure is an important and urgent task. A solution to this problem is proposed in this article.

The existing methods of optimizing the policy of preventive maintenance and repair of technical systems are based on such concepts as “mean time between failures”, “failure rate”, “availability factor”. The solutions based on these characteristics are presented in such well-known works as [1,2,3,4]. Estimations of these characteristics are obtained by processing statistical information from the observation of a wide range of similar systems. It is clear that such estimations can characterize only a generic system and cannot take into account the specifics of a current condition of a specific system. In addition, it is problematic to obtain a representative statistical sample of a sufficiently complex system.

There are some definitions of condition of a complex system which have two values and do not take into account the structure [5]. Other definitions can, for example, introduce a logical variable for identifying the subsystem condition or divide the system operating time into three stages: “burn-in”, “normal work”, “wear-out” and assign the condition to each stage. But there is a difficulty in

identifying the condition of a working system. In any case, only generic solutions are available to maintain reliability at a given level, which, obviously, could not be high enough.

A possible option of a system condition is proposed in [6, 7]. The condition in the model is introduced as the value of a random process, the realizations of which are functionally related to the failure probability. The dependence of the failure probability on the condition is estimated according to statistical data. For the convenience of obtaining this dependence, a random process with monotonically increasing trajectories is chosen. All these actions are important for constructing not a generic but specific model. The disadvantage of this approach is that it is impossible to properly define the system condition at the moment of inspection and therefore to make an effective decision.

The model proposed below is a condition-based one. When determining the condition, the following requirements have been accounted for. The assessment of specific system condition in accordance with the observed values of the parameters as well as the available a priori information for each subsystem should be available at any inspection time. It should be noted that the construction of a condition-based mathematical model of system reliability are described as a promising one in [1]. The proposed condition-based model of a system is more suitable for solving the problem of preventive maintenance of a complex system than the ones mentioned above in a brief review of existing approaches, since it takes into account the specific features of a system under consideration. But this method requires larger amount of statistical information.

The system condition forms the basis for the mathematical model of a complex system. The system condition should include the degree of wear of each subsystem, the failure probability of each subsystem, and the interconnection of subsystems in terms of reliability. System failure is allowed. Therefore, reducing the probability of system failure is one of the main tasks for improving the efficiency of the system. It is assumed that at the moment of system failure one of the possible actions is applied. The chosen action ensures that all failed subsystems are restored. Additionally, the chosen action determines one of the subsystems to be updated reducing the likelihood of its failure. We assume that the duration of an emergency shutdown depends on the set of failed subsystems that has caused the failure of a complex system. If there is no failure, the system is stopped for maintenance after operating for a time interval  $\tau$  specified by the regulations. The duration of the stop is specified by the regulations as well. One of possible actions is applied to ensure the restoration of all failed subsystems to an operating condition, and one of the subsystems determined by the selected action is updated.

When constructing a model, a procedure for determining the basic sets of subsystems the failure of which causes the complex system failure is proposed. We assume that the cost of restoring a system depends on the specified set and the costs of restoring its subsystems, as well as the costs of actions and downtimes are known.

The model is supposed to use certain statistics. All of them refer to the operation of subsystems on time intervals of duration  $\tau$ . Statistical requirements and their usage will be considered during the model construction process. Real statistical information may be incomplete; therefore it is allowed to use expert assessments for the model construction. The system operation being observed, the missing information can be supplemented and the model refined.

The constructed model allows us to determine the optimal policy for choosing actions at planned and unplanned moments of inspection (action policy). This means that at each moment of inspection all failed subsystems will be restored, and the subsystem for updating will be selected. These actions provide an optimal level of reliability and system performance for an unlimited time interval under the specified conditions and constraints.

The proposed work considers a complex system consisted of simple subsystems. The analyst determines whether the system is simple or complex taking into account the purpose of modeling, the chosen modeling methods, available information about the system, and the required degree of adequacy. To a large extent, the adequacy of the model of a complex system depends not only on the available information, but on the analyst's skill as well.

## 2 Mathematical model

The purpose of constructing a mathematical model is to optimize the reliability and performance of a complex technical system operating for an unlimited time interval. The model will be based on the system condition. The condition can be assessed at every moment of inspection. It characterizes the performance of the system and the failure probability. System failure as a result of wear of subsystems

during operation is allowed as well as its restoration to working condition. Preventive maintenance and repair are supposed to be performed at the specified time intervals or in case of failure. The paper proposes a method for optimizing the choice of preventive maintenance and repair, which considers the system condition.

In this work the model of a complex system built in [8] is modified, the statistical requirements is changed, the assessment of the current system condition is simplified and refined, the adequacy of the model is increased, and the class of modeled systems is expanded.

The model is constructed in terms of the Markov decision-making process [9,10]. The method based on the Banach contraction principle is used for the optimization [9].

Let a complex technical system  $S$  be composed of a finite number  $n$  of simple subsystems and be examined on an unlimited time interval. It is required that at any moment of time it is possible to unambiguously determine whether the subsystem is in working or non-working (failure) condition. It is convenient to assume that each simple subsystem has one “input” and one “output”. Then the working condition of a simple subsystem can be associated with the presence of an input-output connection, and the non-working condition with the absence of such connection. The model considers and analyzes the sets of several subsystems. Let each set have one input and one output. That will unambiguously identify the working or failure condition of the set in question at any given time.

Simple subsystems are arbitrarily interconnected in terms of reliability. Isolated subsystems are excluded. An example of such complex system is presented in the Appendix.

## 2.1 System reliability function

It is obvious that the reliability of the system  $S$  should depend on the reliability of its constituent subsystems. Let us consider some properties of separate subsystems. Let us adopt the following definition of the reliability function. Let  $\xi_i$  be the random time before the failure of the  $i$ -th subsystem,  $i=1, \dots, n$ . Then  $P^{(i)}(t) = P(\xi_i \geq t)$  is a reliability function of the  $i$ -th subsystem [1].

Each subsystem, as a rule, has a set of monitored parameters containing information about its reliability and performance. Let the integral indicator of reliability determined by these parameters be denoted by  $\theta^{(i)}$ ,  $i=1, \dots, n$ . Let the parameter  $\theta^{(i)}$  increase monotonically with increasing wear of the subsystem on the interval  $\Theta^{(i)} = [\underline{\theta}^{(i)}, \bar{\theta}^{(i)}]$  [11].

We assume that the system  $S$  is inspected at regular time intervals  $\tau$  if there is no system failure and at the moment of failure if it occurs. At each moment of inspection, one of the possible actions aimed at restoring the failed subsystems and preventive maintenance (update) of one of the subsystems is applied. Let us denote a random variable  $\zeta = \min\{\xi, \tau\}$ , where  $\xi$  the random time before the system failure is,  $\tau$  is the time interval specified by the regulations as the interval of inspection for the system  $S$ . The beginning of any interval is bound to a  $t=0$  moment.

Let us suppose that there is the information about the parameter changes of each of the subsystems at some inspection intervals  $\zeta$ . We assume that it can be limited only to the values at the time of inspection and at the time of subsystem failure. Let us choose the intervals where the information about the trajectory of the parameter of the  $i$ -th subsystem is available. Fig. 1 shows an example of three observed trajectories placed on the same interval  $[0, \tau)$ . At the beginning of the interval at  $t=0$  the parameter values are equal to  $\theta^{(i)}(0)$ . Using interpolation and extrapolation methods, it is possible to estimate the dependence of the parameter of the new subsystem on time [12]. Let us denote it as  $\theta^{(i)}(t)$ ,  $t \in [0, T]$ ,  $\theta^{(i)}(0) = \underline{\theta}^{(i)}$ , where  $T$  is the maximum duration of the time interval provided by statistical data,  $i=1, \dots, n$ . In Fig. 1 it is continuous over the time interval  $[0, T]$ .

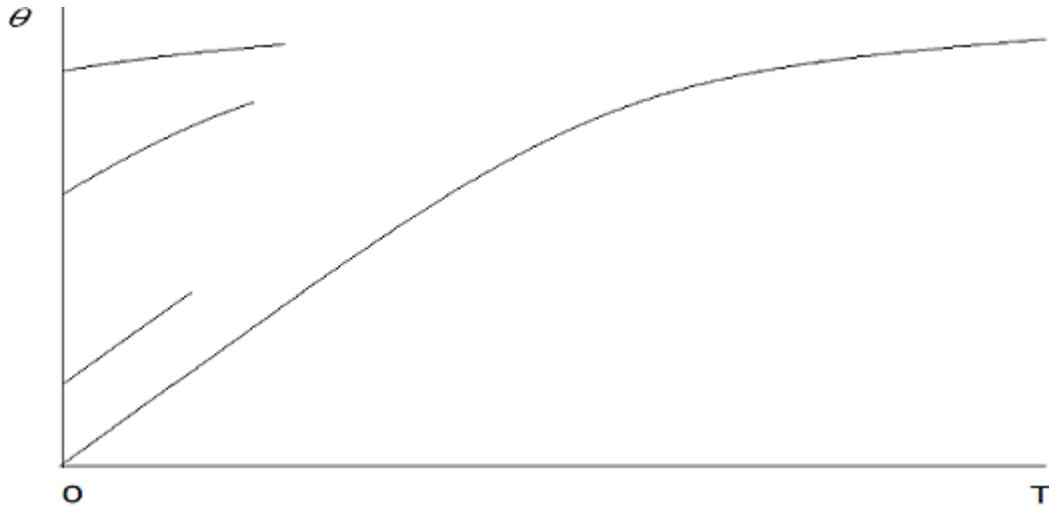


Fig. 1 The estimation of the dependence of the subsystem parameter on time

The adequacy of the model depends on the accuracy of estimating the function  $\theta^{(i)}(t)$ . Therefore, when changing the operating conditions of the system or obtaining new statistical data on the evolution of a parameter, it is recommended to refine the assessment of its  $\theta^{(i)}(t)$  functions.

Let us denote the function inverse to  $\theta^{(i)}(t)$ , as  $t = \psi^{(i)}(\theta)$ ,  $i = 1, \dots, n$ . It should be noted that  $\psi^{(i)}(\underline{\theta}^{(i)}) = 0$ .

**Discretization of the set of parameter values.** To construct the model it is necessary to discretize the set of parameter values for each subsystem of the system  $S$ . For simplicity's sake, we assume that the number of parameter values of any subsystem is equal to  $m$ . Let us denote  $\delta = \frac{\bar{\theta}^{(i)} - \underline{\theta}^{(i)}}{m}$ . Let us assign the discrete parameter value  $\theta_k^{(i)} = \underline{\theta}^{(i)} + (k-1)\delta$ ,  $k = 1, \dots, m-1$  to the half-interval  $[\underline{\theta}^{(i)} + (k-1)\delta, \underline{\theta}^{(i)} + k\delta)$  and the discrete parameter value  $\theta_m^{(i)} = \underline{\theta}^{(i)} + (m-1)\delta$  to the closed interval  $[\underline{\theta}^{(i)} + (m-1)\delta, \bar{\theta}^{(i)}]$ . The discrete set of values of the parameter of the  $i$ -th subsystem will be denoted by  $\tilde{\Theta}^{(i)}$ , and a separate value will be denoted by  $\tilde{\theta}^{(i)}$  or by  $\tilde{\theta}_k^{(i)}$ ,  $k = 1, \dots, m$ ,  $i = 1, \dots, n$ .

Let the  $i$ -th subsystem have the parameter value  $\theta^{(i)} \in \Theta^{(i)}$  at the moment of inspection. It corresponds to a discrete parameter  $\tilde{\theta}_j^{(i)} \in \tilde{\Theta}^{(i)}$ , the value of which can be calculated by the formula

$$\tilde{\theta}_j^{(i)} = \sum_{k=1}^{m-1} h^+ \left( \theta^{(i)} - (\underline{\theta}^{(i)} + (k-1)\delta) \right) h^- \left( \underline{\theta}^{(i)} + k\delta - \theta^{(i)} \right) (\underline{\theta}^{(i)} + (k-1)\delta) + h^+ \left( \theta^{(i)} - (\underline{\theta}^{(i)} + (m-1)\delta) \right) h^+ \left( \underline{\theta}^{(i)} + m\delta - \theta^{(i)} \right) (\underline{\theta}^{(i)} + (m-1)\delta),$$

where  $h^+(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}$ ,  $h^-(t) = \begin{cases} 1, & t > 0 \\ 0, & t \leq 0 \end{cases}$  – the variants of the Heaviside function.

**Actions.** Let a set of actions  $Y = Y_1 \cup Y_2$  be given. Any action in the set  $Y_1 = \{y^{(1)}, \dots, y^{(n)}\}$  can be applied at the scheduled time of inspection. Any action from the set  $Y_2 = \{\hat{y}^{(1)}, \dots, \hat{y}^{(n)}\}$  can be applied at the moment of system failure. The cost of actions depends on the subsystem selected for updating.

Let us consider the effect of action  $y^{(i)} \in Y_1$ . At the scheduled time of inspection, the application of action  $y^{(i)}$  leads to the restoration of all failed subsystems and to preventive maintenance of the  $i$ -th subsystem. Assume that the following values of the subsystem parameters are recorded after the restoration of the failed subsystems:  $\tilde{\theta}^{(j)} \in \tilde{\Theta}^{(j)}$ ,  $j=1, \dots, n$ . Application of action  $y^{(i)}$  changes the value of the parameter of the  $i$ -th subsystem to the value  $\tilde{\theta}_k^{(i)} \in \tilde{\Theta}^{(i)}$  with probability  $\rho\left(\tilde{\theta}_k^{(i)} \mid \tilde{\theta}^{(i)}, y^{(i)}\right)$ ,  $k=1, \dots, m$ . The estimation of the probability distribution  $\rho$  can be obtained on the basis of statistical data or by using expert estimations. The distribution  $\rho$  for actions from  $Y_2$  can be the same as for actions from  $Y_1$  or it can be estimated separately.

**Subsystem reliability function.** Numerical information about the probability of subsystem failure is determined by its reliability function. The estimation of the subsystem reliability function can be obtained after processing the available information about the parameter change over time, as well as subsystem failures. Let us consider the case when a failure of any subsystem is detected at the moment of the failure.

Let  $L$  be the number of intervals that contain available information. At the beginning of the  $l$ -th interval, the action having been applied, the value of the parameter  $\theta^{(i)}(l) \in \Theta^{(i)}$  of the  $i$ -th subsystem,  $i=1, \dots, n$  is recorded.

If the  $i$ -th subsystem does not fail on the  $l$ -th interval, then at the time of inspection  $\tau$  the subsystem is in working condition and the parameter value  $\theta^{(i)}(l, \tau) \in \Theta^{(i)}$  is recorded. If the  $i$ -th subsystem on the  $l$ -th interval fails at a  $t^{(i)}(l, \xi)$  moment, then the value of the parameter  $\theta^{(i)}(l, \xi) \in \Theta^{(i)}$  is recorded.

We convert the available statistical information on the  $i$ -th subsystem for the new  $i$ -th subsystem in the following way. Let us set the  $l$ -th interval. If at the scheduled time of inspection  $\tau$  the parameter value equals  $\theta^{(i)}(l, \tau) \in \Theta^{(i)}$ , then for a new subsystem this value is reached at the time  $\tilde{t}^{(i)}(l, \tau) = \psi^{(i)}\left(\theta^{(i)}(l, \tau)\right)$ . If at the moment of failure the parameter value equals  $\theta^{(i)}(l, \xi) \in \Theta^{(i)}$ , then in a new subsystem the failure would occur at the moment of time  $\tilde{t}^{(i)}(l, \xi) = \psi^{(i)}\left(\theta^{(i)}(l, \xi)\right)$ . Note that if at the moment of failure the parameter value is unknown, but the time of failure  $t^{(i)}(l, \xi)$  is known, then  $\tilde{t}^{(i)}(l, \xi) = t^{(i)}(l, \xi) + \psi^{(i)}\left(\theta^{(i)}(l)\right)$ , where  $\theta^{(i)}(l)$  is the parameter value at the beginning of the  $l$ -th interval. Let the shifted time moments  $\tilde{t}^{(i)}(l, \tau)$ ,  $\tilde{t}^{(i)}(l, \xi)$  be found for all intervals  $l=1, \dots, L$  and for all subsystems  $i=1, \dots, n$ . We denote by  $T^{(i)}$  the maximum value of all  $\tilde{t}^{(i)}(l, \tau)$  and  $\tilde{t}^{(i)}(l, \xi)$  for each  $i$ .

The empirical reliability function of the  $i$ -th subsystem can be defined as follows.

Let us denote by  $L^{(i)}(\xi)$  the set of intervals in which the failure of the  $i$ -th subsystem occurs.

We denote by  $H^{(i)}(t) = \frac{1}{L} \sum_{l \in L^{(i)}(\xi)} h^-\left(t - \tilde{t}^{(i)}(l, \xi)\right)$  if  $L^{(i)}(\xi) \neq \emptyset$  and assume  $H^{(i)}(t) = 0$  if

$L^{(i)}(\xi) = \emptyset$ . Then for time until the failure of the new  $i$ -th subsystem the empirical distribution function has the following form

$$F_0^{(i)}(t) = H^{(i)}(t) + \left(1 - H^{(i)}(t)\right) h^-\left(t - T^{(i)}\right),$$

and the empirical reliability function has the form

$$P_0^{(i)}(t) = 1 - F_0^{(i)}(t). \quad (2.1)$$

Afterwards, it is more convenient to use leveled versions of the functions  $F_0^{(i)}(t)$  and  $P_0^{(i)}(t)$ , where  $F_0^{(i)}(0) = 0$  and at  $t \rightarrow \infty$   $F_0^{(i)}(t)$  monotonously tends to 1.

The reliability functions of the new subsystems  $P_0^{(i)}(t)$  define the reliability function  $P_0(t)$  of the new system  $S$ . The reliability function for the model of the system is given in Appendix Item 1.

**The prediction of the reliability function.** The prediction of the reliability function of the new  $i$ -th subsystem at the moment of time  $\delta$  can be estimated by the function

$$P_\delta^{(i)}(t) = \frac{P_0^{(i)}(t + \delta)}{P_0^{(i)}(\delta)}, \quad t > 0, \quad i = 1, \dots, n. \quad (2.2)$$

Here  $P_0^{(i)}(t)$ ,  $i = 1, \dots, n$ , the reliability functions of new subsystems. They define the reliability function  $P_\delta(t)$  of the new system  $S$  at a point in time  $\delta$ .

Assume that all failed subsystems having been restored, the parameters of the subsystems take on the values  $\theta^{(i)}$ ,  $i = 1, \dots, n$  at the time of inspection  $t_c$ . For each  $i$ -th subsystem, we will find the shift of values  $\hat{t}^{(i)} = \psi^{(i)}(\theta^{(i)})$ .

The estimation of the reliability function of the  $i$ -th subsystem at the time of inspection  $t_c$  has the form

$$P_c^{(i)}(t) = \frac{P_0^{(i)}(t + \hat{t}^{(i)})}{P_0^{(i)}(\hat{t}^{(i)})}, \quad t > 0, \quad i = 1, \dots, n. \quad (2.3)$$

Then the system reliability function at the time of inspection  $t_c$  can be determined with the functions  $P_c^{(i)}(t)$ ,  $i = 1, \dots, n$ . We denote it by  $P_c(t)$ .

The reliability function of the  $i$ -th subsystem in a time interval  $\delta$ , after evaluating the parameters at a time moment  $t_c$ , can be represented in the form

$$P_\delta^{(i)}(t) = \frac{P_0^{(i)}(t + \hat{t}^{(i)} + \delta)}{P_0^{(i)}(\hat{t}^{(i)} + \delta)}, \quad t > 0, \quad i = 1, \dots, n. \quad (2.4)$$

Reliability functions  $P_\delta^{(i)}(t)$ ,  $i = 1, \dots, n$ , determine the system  $S$  reliability function  $P_\delta(t)$  in the time interval  $\delta$  after the moment of inspection.

## 2.2 System condition

Let the condition  $x$  of the new system  $S$  at the initial moment of time be equal to 0.

We can calculate the average operating time  $\bar{t}$  of the new system  $S$  over a time interval  $\tau$  using the Stieltjes integral

$$\bar{t} = \int_0^\tau t d(1 - P_0(t)) + \tau \cdot P_0(\tau). \quad (2.5)$$

Let  $P_c(t)$  be the reliability function of the system  $S$  at the time of inspection  $t_c$ . Then the average operating time  $\bar{t}_c$  of the system over a time interval  $(t_c, t_c + \tau)$  is

$$\bar{t}_c = \int_0^{\tau} t d(1 - P_c(t)) + \tau \cdot P_c(\tau). \quad (2.6)$$

Let us assume that at the moment of inspection  $t_c$  the condition of the system  $S$  is equal to  $x_c = \frac{\bar{t} - \bar{t}_c}{\bar{t}}$ .

The prediction of the reliability function  $P_\delta(t)$  of the system  $S$  in the time interval  $\delta$  after the moment of inspection  $t_c$  allows us to calculate

$$\bar{t}_\delta = \int_0^{\tau} t d(1 - P_\delta(t)) + \tau \cdot P_\delta(\tau). \quad (2.7)$$

Let us assume that the condition of the system  $S$  at time  $t_c + \delta$  is equal to  $x_\delta = \frac{\bar{t} - \bar{t}_\delta}{\bar{t}}$ .

As a result, the condition of the new system  $S$  at the initial moment of time is  $x=0$ , and the condition of a completely worn out system is close to 1. The set  $E=[0,1]$  contains all possible conditions.

In addition, it has been found that the set of parameters of the subsystems  $\{\theta^{(i)}, i=1, \dots, n\}$ , recorded at the moment of inspection  $t_c$  after applying the action, makes it possible to estimate the condition of the system  $S$  at any moment  $t$ ,  $t_c \leq t \leq \tau$ . It is assumed that estimations of the reliability function of all subsystems have been obtained earlier.

**Discretization of the set of conditions.** Let us suppose that the discrete set of conditions of the system  $S$  consists of  $N$  elements and denote it by  $\tilde{E}$ . Let us assign the discrete condition  $x_k$  of the system  $S$ ,  $k=0, \dots, N-1$  to the half-interval  $\left[ k \frac{1}{N}, (k+1) \frac{1}{N} \right)$ . Let us assume that the condition  $x_k$  has the value  $\frac{k}{N}$ .

Let the condition  $x \in E$  be observed at the moment of inspection of the system  $S$ . It corresponds to the discrete condition  $\tilde{x}_v$ , the value of which can be found by the formula

$$\tilde{x}_v = \Lambda(x) = \sum_{k=0}^{N-2} h^- \left( \frac{k+1}{N} - x \right) h^+ \left( x - \frac{k}{N} \right) \frac{k}{N} + h^+ (1-x) h^+ \left( x - \frac{N-1}{N} \right) \frac{N-1}{N}. \quad (2.8)$$

Here  $\Lambda$  is a mapping of a set  $E$  to a discrete set of conditions  $\hat{E} = (x_0, \dots, x_{N-1})$ .

### 2.3 The process of changing the system condition over time

Subsystems wear out during operation, and some of them may fail at random times. Information about possible failures of subsystems is accounted in the corresponding reliability functions. In addition, the presence of unaccounted factors determines the evolution of the system condition over time as a random process.

If at the beginning of the interval all subsystems are new, then the reliability functions of new subsystems at the moment of time  $\delta$  are presented in (2.2). Let us consider these functions, assuming  $\delta$  to be a variable. The corresponding reliability function of the new system  $S$  will depend on the variable  $\delta$ , which we will further denote by  $z$ . The time-dependence function of the new system



condition will be  $x(z) = \frac{\bar{t} - \bar{t}(z)}{\bar{t}} \in E$ , where  $\bar{t}(z) = \int_0^z t d(1 - P_z(t)) + \tau P_z(\tau)$ . An explicit form of the dependence and its graph for a specific system can be obtained in the MAPLE system (see Appendix Item 2.).

Let us assume that the dependence of the trajectory of the condition  $x(t), t \geq 0$ , on time has the form

$$x(t) = \hat{x}(0) + (1 - \hat{x}(0)) \frac{\gamma \cdot t}{\gamma \cdot t + \beta}, \quad t \geq 0, \quad x(t) \in E,$$

where  $\hat{x}(0)$  is the initial condition.

The estimations of  $\gamma$  and  $\beta$  parameters can be obtained from observations of a single trajectory (for example, a new system for which  $\hat{x}(0) = 0$ ). If necessary, the shape of the trajectory can be corrected by observing the trajectories of a specific system.

Suppose that at the scheduled time of inspection  $t_c = 0$ , after the restoration of all failed subsystems, a condition  $\hat{x}(0) \in \hat{E}$  is observed and the values of the parameters  $\{\theta^{(j)}, j = 1, \dots, n\}$  is recorded. The application of action  $y^{(i)} \in Y$  will determine the sets of parameter values  $A_k = \{\theta^{(1)}, \dots, \theta_k^{(i)}, \dots, \theta^{(n)}\}$ ,  $k = 1, \dots, m$ . Each set  $A_k$  determines the system condition  $x(0, k) \in E$  at the beginning of the interval. Let us denote  $\hat{x}(0, k) = \Lambda(x(0, k)) \in \hat{E}_1$ . For each  $j$ ,  $j = 0, \dots, N-1$ , we define the set  $K_j = \{k : \Lambda(x(0, k)) = \hat{x}_j(0)\}$ , where  $\hat{x}_j(0)$  is the condition with the number  $j$  from the set  $\hat{E}$  at the moment of time  $t_c = 0$ . Let us assume that the desired probability of the discrete condition  $\hat{x}_j(0)$  after applying the action  $y^{(i)} \in Y$  is

$$P(\hat{x}_j(0) | \hat{x}(0), y^{(i)}) = \sum_{k \in K_j} \rho(\theta_k^{(i)} | \theta^{(i)}, y^{(i)}), \quad j = 0, \dots, N-1. \quad (2.9)$$

**Transition probability.** For further model construction, it is convenient to expand the discrete set of conditions  $\hat{E}$ . If a system fails in the condition  $\hat{x}_j \in \hat{E}$ , then we will formally consider it a different condition and denote it by  $\tilde{x}_j$ . In this case, the set of conditions will be doubled. We denote the resulting discrete set of conditions by  $\tilde{E} = \hat{E} \cup \tilde{E}' = \{\hat{x}_0, \dots, \hat{x}_{N-1}, \tilde{x}_0, \dots, \tilde{x}_{N-1}\}$ . We assume that in the conditions of the set  $\hat{E}$ , the actions from the set  $Y_1$ , and in the conditions of the set  $\tilde{E}'$ , the actions from the set  $Y_2$  can be applied. The sets  $Y_1$  and  $Y_2$  do not intersect and  $Y = Y_1 \cup Y_2$ .

Let us calculate the probability of each condition from  $\tilde{E}$  which the system can assume in a time interval  $z$  after applying the action  $y^{(i)} \in Y$  in a given condition from the set  $\tilde{E}$ .

If the action  $y^{(i)} \in Y_1$  is applied in the condition  $\hat{x}(0) \in \hat{E}$ , then the probability of the condition  $\hat{x}_j(z)$  that the system will have in the time interval  $z$  is denoted by  $p(\hat{x}_j(z) | \hat{x}(0), y^{(i)})$ . It should be noted that after applying the action  $y^{(i)} \in Y_1$  in the condition  $\hat{x}(0) \in \hat{E}$ , the system transitions to the condition  $\hat{x}_j(0)$  with probability  $P(\hat{x}_j(0) | \hat{x}(0), y^{(i)})$ ,  $j = 0, \dots, N-1$ .

Let us denote  $K_j(z) = \left\{ k : \Lambda\left(\hat{x}_k(0) + (1 - \hat{x}_k(0)) \frac{\gamma z}{\gamma z + \beta}\right) = \hat{x}_j \right\}$ .

Then

$$P\left(\hat{x}_j(z) \mid \hat{x}(0), y^{(i)}\right) = \sum_{k \in K_j(z)} p\left(\hat{x}_k(0) \mid \hat{x}(0), y^{(i)}\right), \quad z \in [0, \tau]. \quad (2.10)$$

If the time of inspection  $t_c = 0$  is unplanned, then the form of formula (2.10) remains correct, provided that  $\hat{x}(0) \in \hat{E}$  is replaced by  $\tilde{x}(0) \in \tilde{E}'$  and  $y^{(i)} \in Y_1$  by  $y^{(i)} \in Y_2$ .

Let the trajectory of the system condition after the application of the action start from  $\hat{x}_j(0)$ . Then

$$\hat{x}_j(0) = \frac{\bar{t} - \hat{t}_{md}(\chi_j)}{\bar{t}}, \quad \text{where } \bar{t} \text{ is defined in (2.5), } \hat{t}_{md}(\chi_j) = \int_0^{\tau} t \cdot d \left( 1 - \frac{P_0(t + \chi_j)}{P_0(\chi_j)} \right) + \tau \frac{P_0(\tau + \chi_j)}{P_0(\chi_j)}.$$

Since  $\hat{t}_{md} = \bar{t}(1 - \hat{x}_j(0))$ , the value of the shift  $\chi_j$  of the reliability function of the new system for the

initial condition  $\hat{x}_j(0) = \frac{j}{N}$  is found from the equation  $\hat{t}_{md}(\chi_j) = \bar{t} \left( \frac{N-j}{N} \right)$ ,  $j = 0, \dots, N-1$ . The

shifts  $\chi_j$ ,  $j = 0, \dots, N-1$  can be calculated in the MAPLE system for all conditions  $\hat{E}$  during the model construction before its optimization. These values can be used to construct transition and direct income functions.

#### 2.4 Decision function. Transition matrix

The transition function  $Q\left(x_v \mid x_u, y^{(i)}, t\right)$ ,  $x_v \in \hat{E}$ , determines the probability distribution on the set  $\tilde{E}$  at the moment of the next inspection (planned at  $t = \tau$  or unplanned at the moment of system failure), provided that  $x_u \in \tilde{E}$  is the condition of the system in which action  $y^{(i)} \in Y$  is applied at the beginning of the interval.

Let the mapping  $f: \tilde{E} \rightarrow Y$  be given, and  $f(\hat{x}_j) \in Y_1$ ,  $f(\tilde{x}_j) \in Y_2$ ,  $j = 0, \dots, N-1$ . The mapping  $f$  is called a decision function. The sequence of decision functions  $\pi = \{f_0, f_1, \dots\}$  is called a action policy. The policy  $\pi = f^{(\infty)} = \{f, f, \dots\}$  is called stationary [9].

For each decision function  $f$  there is a corresponding matrix of transition probabilities  $Q(f)$  of dimension  $2N \times 2N$ .

$$Q(f) = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix},$$

We will assume that the matrix  $Q(f)$  is regular for any decision function  $f$  [9]. In our case, this limitation is not essential.

Let us introduce the notation for the elements of the submatrices  $Q_{ij}$ :

$$Q_{11} = \begin{pmatrix} q(\hat{x}_0 \mid \hat{x}_0, f(\hat{x}_0)) & \dots & q(\hat{x}_{N-1} \mid \hat{x}_0, f(\hat{x}_0)) \\ \dots & \dots & \dots \\ q(\hat{x}_0 \mid \hat{x}_{N-1}, f(\hat{x}_{N-1})) & \dots & q(\hat{x}_{N-1} \mid \hat{x}_{N-1}, f(\hat{x}_{N-1})) \end{pmatrix},$$

$$Q_{12} = \begin{pmatrix} q(\tilde{x}_0 \mid \hat{x}_0, f(\hat{x}_0)) & \dots & q(\tilde{x}_{N-1} \mid \hat{x}_0, f(\hat{x}_0)) \\ \dots & \dots & \dots \\ q(\tilde{x}_0 \mid \hat{x}_{N-1}, f(\hat{x}_{N-1})) & \dots & q(\tilde{x}_{N-1} \mid \hat{x}_{N-1}, f(\hat{x}_{N-1})) \end{pmatrix},$$

$$Q_{21} = \begin{pmatrix} q(\hat{x}_0 | \tilde{x}_0, f(\tilde{x}_0)) & \dots & q(\hat{x}_{N-1} | \tilde{x}_0, f(\tilde{x}_0)) \\ \dots & \dots & \dots \\ q(\hat{x}_0 | \tilde{x}_{N-1}, f(\tilde{x}_{N-1})) & \dots & q(\hat{x}_{N-1} | \tilde{x}_{N-1}, f(\tilde{x}_{N-1})) \end{pmatrix},$$

$$Q_{22} = \begin{pmatrix} q(\tilde{x}_0 | \tilde{x}_0, f(\tilde{x}_0)) & \dots & q(\tilde{x}_{N-1} | \tilde{x}_0, f(\tilde{x}_0)) \\ \dots & \dots & \dots \\ q(\tilde{x}_0 | \tilde{x}_{N-1}, f(\tilde{x}_{N-1})) & \dots & q(\tilde{x}_{N-1} | \tilde{x}_{N-1}, f(\tilde{x}_{N-1})) \end{pmatrix}.$$

The matrix element  $Q_{11}$  can be calculated with allowance for (2.9) and (2.10) by the formula

$$q(\hat{x}_j | \hat{x}_k, y) = \sum_{\kappa \in K_j(\tau)} p(\hat{x}_\kappa(0) | \hat{x}_k, y) \frac{P_0(\tau + \chi_\kappa)}{P_0(\chi_\kappa)}, \quad (2.11)$$

where  $K_j(\tau) = \left\{ k : \Lambda \left( \hat{x}_k(0) + (1 - \hat{x}_k(0)) \frac{\gamma\tau}{\gamma\tau + \beta} \right) = \hat{x}_j \right\}$ ,  $y = f(\hat{x}_k) \in Y_1$ ,  $\chi_\kappa$  is the shift of the trajectory of new system conditions for the trajectory beginning  $\hat{x}_\kappa(0)$ .

Formula (2.11) remains correct for the matrix elements  $Q_{21}$  after  $\hat{x}_k \in \hat{E}$  is replaced by  $\tilde{x}_k \in \tilde{E}'$  and if  $y = f(\tilde{x}_k) \in Y_2$ .

The matrix elements  $Q_{12}$  can be calculated as follows

$$q(\tilde{x}_j | \hat{x}_k, y) = \int_0^\tau \sum_{\kappa \in K_j(t)} p(\tilde{x}_\kappa | \hat{x}_k, y) d \left( 1 - \frac{P_0(t + \chi_\kappa)}{P_0(\chi_\kappa)} \right), \quad (2.12)$$

where  $K_j(t) = \left\{ \kappa : \Lambda \left( \hat{x}_\kappa(0) + (1 - \hat{x}_\kappa(0)) \frac{\gamma t}{\gamma t + \beta} \right) = \hat{x}_j \right\}$ ,  $y = f(\hat{x}_k) \in Y_1$ ,  $\chi_\kappa$  is the shift of the trajectory of the new system conditions for the trajectory beginning  $\hat{x}_\kappa(0)$ .

Formula (2.12) is correct for the matrix elements  $Q_{22}$  after  $\hat{x}_k$  is replaced by  $\tilde{x}_k \in \tilde{E}'$  and if  $y = f(\tilde{x}_k) \in Y_2$ .

## 2.5 Basic set of failures

Complex system failure can occur as a result of failure of all subsystems in the system  $S$  or one specific set of subsystems. The duration of system downtime, as well as the cost of its restoring to an operational condition usually depends on the set of failed subsystems. Therefore, it is important to define all such sets in advance. We will assume that failure of any simple subsystem does not affect the failure probability of any other simple subsystem.

Simple subsystems wear out during operation and their failure probability increases. Failure of one or more subsystems does not necessarily result in the failure of the system  $S$ . However, there is a number  $M$  of sets of subsystems with the following property. Failure of all subsystems in a set from  $M$  leads to failure of the system  $S$ . None of the set from  $M$  contains subsystems "indifferent" to failure of the system  $S$ : failure of the system  $S$  occurs when all subsystems of the set fail, and after restoring any single subsystem of this set, the system  $S$  will not remain in a failure condition. If  $M$  contains all the different sets of subsystems from  $S$  that have the properties indicated above, then this set will be called a base set of sets leading to failure of the system  $S$ , or simply a basic set of failures. Note that sets of  $M$  can overlap.

**Algorithm for compiling a basic set of failures.** The method for finding the basic set of failures  $M$  which can be easily implemented by using the MAPLE mathematical program is proposed. Let the

system  $S$  consist of  $n$  simple subsystems. The system reliability function is supposed to be known. Let us execute the following algorithm.

1. Find all subsets of a set of  $n$  elements. Denote it by  $C$ . Formally identify an element of the set  $C$  with the corresponding set of subsystems.
2. Form a set  $G$ , consisting of the elements from  $C$ , which correspond to a set of subsystems, the failure of which leads to failure of the system  $S$ . Denote by  $G_i$ , elements of the set  $G$ , which correspond to sets of  $i$  subsystems,  $i=1, \dots, n$ .

Note that, for some  $i$  the set  $G_i = \emptyset$ .

3. If  $G_1 \neq \emptyset$ , then  $\tilde{G}_1 = G_1$ .
4. If  $G_1 = G_2 = \dots = G_l = \emptyset$ ,  $G_{l+1} \neq \emptyset$ , then  $\tilde{G}_{l+1} = G_{l+1}$ .
5. Let  $G_k \neq \emptyset$ . Define the set  $\tilde{G}_k$ . Any subset  $g \in G_k$  is excluded from  $G_k$  if there is a subset  $h \in \tilde{G}_{k-j}$ ,  $j=1, \dots, k-1$ , such that  $g \supset h$ . The remaining elements form a set  $\tilde{G}_k$ . The union of the obtained non-empty sets  $\tilde{G}_i$  forms the basic set of failures  $M$ .

The example of implementing this algorithm for the model can be found in Appendix Item 3.

The theoretical calculation of the failure probability of the given set  $h \in M$  up to the moment of time  $t$  that leads to the failure of the system  $S$  is not difficult. However, applying this data to the model would make determining the condition rather cumbersome. The practical value of the model will be significantly lower due to the complexity of the implementation. Therefore, the method for accounting the sets of failures from  $M$  is proposed.

Suppose that failure of the system  $S$  occurs during a interval of duration  $\tau$ . In this case, the only set from  $M$  is realized. Let us denote the cost of restoring the system  $S$  to a working condition by  $g(h)$ ,  $h \in M$ , if the failure occurred as a result of failure of all subsystems in the set  $h \in M$ . Let us denote the number of elements in  $M$  by  $J$ . A selective distribution of the cost of restoring a failed system can be obtained from the available observations. Therefore, only intervals with failures and the beginnings of trajectories should be taken into account. The probability that the cost of restoration will be  $g_j$ ,  $j=1, \dots, J$ , if the trajectory of the process starts in the condition  $\hat{x}(0) \in \hat{E}$  and the system failure occurs is denoted by  $p(g_j | \hat{x}(0))$ .

## 2.6 Direct income function

Let us denote the average cost of restoring failed subsystems at the time of inspection by  $R$ ; the cost of the action  $f(x) \in Y$  determined by the decisive function  $f$  in the condition  $x \in \tilde{E}$  by  $r(f(x))$ ; the income per unit of time that the operating system yields for one interval starting in the condition  $x \in \tilde{E}$  by  $v(x)$ ; the average duration of system downtime during preventive maintenance by  $\delta$ ; the average duration of system downtime when restoring from a failure condition by  $\Delta$ ; the cost of restoring the system from a failure condition if the failure occurs due to the realization of the set  $h_j \in M$  by  $g_j$ .

The direct income function  $w(x, f(x))$ ,  $x \in \tilde{E}$ ,  $f(x) \in Y$ , with the fixed decision function  $f$  is a column vector composed of  $2N$  components. The component of this vector for a given  $x \in \tilde{E}$  can be calculated by the formula

$$w(x, f(x)) = \sum_{k=0}^{N-1} \frac{v(\hat{x}_k) \cdot \tau}{\tau + \delta} \left( p(\hat{x}_k | x, f(x)) - R - r(f(x)) \right) \frac{P_0(\tau + \chi_k)}{P_0(\chi_k)} + \\ + \int \sum_{k=0}^{\tau} \sum_{j=1}^J \frac{v(\hat{x}_k) \cdot t}{t + \Delta} \left( p(\hat{x}_k | x, f(x)) - R' - r(f(x)) - g_j \cdot p(g_j | \hat{x}_k) \right) d \left( 1 - \frac{P_0(t + \chi_k)}{P_0(\chi_k)} \right).$$

## 2.7 Optimality criterion. Optimization method

We assume that the inequality  $a \geq b$  is true for the vectors  $a$  and  $b$  if a similar inequality  $a_i \geq b_i$  is true for all components  $a_i \geq b_i$ . The inequality  $a > b$  is satisfied if  $a \geq b$  and  $a \neq b$ .

The case of an unlimited system operation time is being considered. The action policy is the selection of admissible actions at the moments of inspection. For each stationary policy  $\pi = f^{(\infty)} = \{f, f, \dots\}$  we identify the corresponding average income per unit of time  $\phi(\pi)$  on an unlimited time interval [9]:

$$\phi(\pi) = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=0}^{L-1} Q^l(f)w(f).$$

The limit exists and the column vector  $\phi(\pi)$  is composed of the components of equal size if the matrix  $Q(f)$  is regular.

A stationary policy  $\pi^*$  is optimal if the inequality  $\phi(\pi^*) \geq \phi(\pi)$  is true for any stationary policy  $\pi$ .

A stationary policy  $\pi_\varepsilon^*$  is  $\varepsilon$ -optimal if the inequality  $\phi(\pi) - \phi(\pi_\varepsilon^*) \leq \varepsilon$  is true for a given  $\varepsilon > 0$  and any stationary policy  $\pi$ .

The search for an  $\varepsilon$ -optimal policy for controlling the reliability of a complex system can be implemented by a well-known method based on the Banach contraction principle [9]. Let us introduce the following auxiliary definitions and notations.

Let a seminorm  $\tilde{p}(v) = \max_i v_i - \min_i v_i$ ,  $1 \leq i \leq 2N$  be defined on a vector space  $V$  with elements  $v$  of dimension  $2N$  [13]. The chosen optimization method requires a normalized space. To fulfill this condition we factorize  $V$  by  $K = \{v: p(v) = 0\}$  to determine the factor-space  $V' = V / K$  on which  $\tilde{p}$  is the norm  $\|\cdot\|$  [13].

Let us define the following operators on  $V$ .

$$F(f)v = w(f) + Q(f)v;$$

$$Uv = \max_f F(f)v.$$

Under the assumption that transition matrix is regular for any decision function  $f$ , the operator  $U$  on  $V'$  is contractive [9]. Those assumptions allow us to apply an optimization method based on the Banach contraction principle [9]. The algorithm stopping rule will be based on the following statement:

Statement [14]. Let  $v \in V'$ ,  $\|Uv - v\| = \varepsilon$ ,  $Uv = F(f)v$ . Then

1. The policy  $\pi = f^{(\infty)}$  is  $\varepsilon$ -optimal.
2.  $\min_x (Uv - v)(x) \leq \phi^* \leq \max_x (Uv - v)(x)$ .

Here  $\phi^*$  is an average income per unit of time, which corresponds to an optimal policy.

Corollary [14]. Let  $\pi^* = (f^*, f^*, \dots)$  be the optimal policy.  $v_k, v_{k+1} \in V'$  and  $f_{k+1}$  satisfy the conditions:  $v_{k+1} = Uv_k = F(f_{k+1})v_k$ . Let us denote  $\varepsilon = \|v_{k+1} - v_k\|$ . Then the decisive function  $f_{k+1}$  determines the  $\varepsilon$ -optimal policy  $\pi_\varepsilon^* = (f_{k+1}, f_{k+1}, \dots)$ .

### Optimization algorithm [9].

Select  $\varepsilon > 0$ .

Choose a starting vector  $v_0 \in V'$  arbitrary.

$k$ -th step of the algorithm ( $k = 1, 2, \dots$ ).

1. Calculate  $v_k = Uv_{k-1} = F(f_k)v_{k-1}$ ,  $v_k \in V'$ .
2. Check the inequality  $\|v_k - v_{k-1}\| < \varepsilon$ .

If it is true, then the policy  $\pi = f_k^{(\infty)}$  is  $\varepsilon$ -optimal.

Otherwise, execute the  $(k + 1)$ -th step of the algorithm.

The  $\varepsilon$ -optimal policy will be achieved for any given  $\varepsilon > 0$  in a finite number of steps [9].

Note 1. It is convenient to choose a  $v_k$  with zero first component from the coset of the factor-space  $V'$ . In practice, this means that for any considered vector  $v$ , it is necessary to subtract the value of the first component from all components of the vector  $v$  in order to obtain  $v' \in V'$ .

### 3 Conclusion

The constructed mathematical model of a complex technical system according to its condition provides a higher level of adequacy and, therefore, allows finding a more effective policy for preventive maintenance and repair for an unlimited time interval in comparison with those proposed earlier in similar conditions. There is no universal mathematical model of a complex system capable to solve any problem. The goal of modeling should always be formulated first.

The article deals with a complex technical system composed of a finite number of simple subsystems which are, generally speaking, different. Each of the simple subsystems wears out during operation, and the probability of its failure increases. A failure of one of the subsystems does not affect the failure probability of any other subsystem. Subsystems are arbitrarily interconnected in terms of reliability. Thus, failure of one or more subsystems does not necessarily lead to failure of the entire system. We assume that the failure of the system leads to long downtime, highly expensive and, therefore, is better to be avoided. The work of subsystems as part of one complex system is to achieve a specified objective. To do that, a modeling approach in terms of a Markov decision-making process, and an optimization method based on the Banach contraction principle have been chosen.

To implement this approach using the available statistical information for each subsystem is proposed. If the volume of statistical information is insufficient, it can be supplemented by expert data. Lack of information reduces the level of model adequacy. On the other hand, using all available statistical information is not always justified due to the increasing model complexity. The practical value of the model can be significantly reduced due to its complexity. In the proposed article a compromise between the complexity of the model and its practical convenience has been sought. Some key points of complex system modeling for the given model example are presented in Appendix.

One of the possible directions for further development of the proposed approach to constructing the models for complex system maintenance according to its condition can be revising, clarifying, and generalizing the concept of action in the problem formulation. Modeling a complex technical system, the work of which is influenced by seasonal fluctuations can be another direction. The optimization method for such a system is proposed in [15]. Optimizing the interval of planned system inspection can be another possibility to improve the efficiency of model usage. The solution to this problem is proposed in [14].

### Appendix

1. Let us consider a model example of the complex technical system  $S$  consisting of  $n = 9$  simple subsystems. The subsystems are interconnected in the sense of reliability as shown in Fig. 2. The connections are denoted by a line crossing.

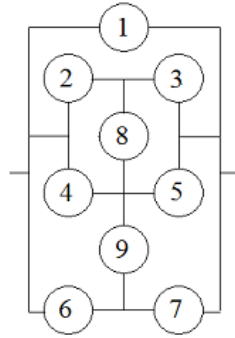


Fig. 2 The complex system example

$P^{(i)}(t)$ ,  $i = 1, \dots, 9$ , are subsystem reliability functions. The method for finding the reliability function  $P(t)$  of the system  $S$  based on the total probability formula has been proposed in [8], namely,

$$P(t) = 1 - (1 - P^{(1)}(t))(1 - PP(t))$$

where:

$$PP(t) = \left[ 1 - (1 - P^{(2)}(t))(1 - P^{(4)}(t)) \right] (1 - P^{(6)}(t)) \left[ 1 - (1 - P^{(3)}(t))(1 - P^{(5)}(t)) \right] \cdot (1 - P^{(7)}(t)) P^{(8)}(t) P^{(9)}(t) +$$

$$+ [1 - (1 - PS_1)(1 - PS_2)] P^{(8)}(t) (1 - P^{(9)}(t)) +$$

$$+ [1 - (1 - PS_3)(1 - PS_4)] (1 - P^{(8)}(t)) P^{(9)}(t) +$$

$$+ [1 - (1 - PS_5)(1 - PS_6)] (1 - P^{(8)}(t)) (1 - P^{(9)}(t)),$$

$$PS_1 = \left[ 1 - (1 - P^{(2)}(t))(1 - P^{(4)}(t)) \right] \left[ 1 - (1 - P^{(3)}(t))(1 - P^{(5)}(t)) \right],$$

$$PS_2 = P^{(6)}(t) P^{(7)}(t),$$

$$PS_3 = P^{(2)}(t) P^{(9)}(t),$$

$$PS_4 = \left[ 1 - (1 - P^{(4)}(t))(1 - P^{(6)}(t)) \right] \left[ 1 - (1 - P^{(5)}(t))(1 - P^{(7)}(t)) \right],$$

$$PS_5 = 1 - (1 - P^{(2)}(t) P^{(3)}(t)) (1 - P^{(4)}(t) P^{(5)}(t)),$$

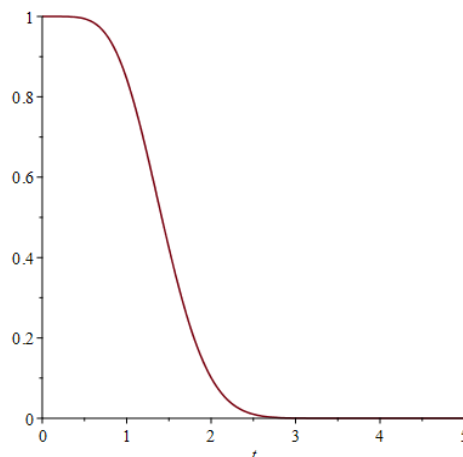


Fig. 3 The new system reliability function

Fig. 3 shows a graph of the reliability function  $p0\_syst(t)$  of the new system  $S$  on the time interval  $[0,5)$

The derivative  $dp0\_syst(t)$  of the reliability function  $p0\_syst(t)$  is found by using the following command:

$$dp0\_syst(t) := diff(p0\_syst(t), t);$$

Let us suppose that  $\tau = 3$ . The average operating time  $T0$  of the new system on the time interval  $[0, \tau)$  can be found with the command:

$$T0 := evalf(int(-t \cdot dp0\_syst(t), t = 0 ..tau) + tau \cdot p0\_syst(tau));$$

$$T0 := 1.437222359 .$$

The prediction of the reliability function  $ptt\_syst(t)$  of the new system at the moment of time  $\delta$  is found by the formula (2). In the MAPLE system, it is implemented with the command ( $\delta$  replaced by  $tt$ ):

$$ptt\_syst(t) := subs\left(p[1] = \left(\frac{p[1](t+tt)}{p[1](tt)}\right), p[2] = \left(\frac{p[2](t+tt)}{p[2](tt)}\right), p[3] = \left(\frac{p[3](t+tt)}{p[3](tt)}\right), p[4] = \left(\frac{p[4](t+tt)}{p[4](tt)}\right), p[5] = \left(\frac{p[5](t+tt)}{p[5](tt)}\right), p[6] = \left(\frac{p[6](t+tt)}{p[6](tt)}\right), p[7] = \left(\frac{p[7](t+tt)}{p[7](tt)}\right), p[8] = \left(\frac{p[8](t+tt)}{p[8](tt)}\right), p[9] = \left(\frac{p[9](t+tt)}{p[9](tt)}\right), p\_syst\right)$$

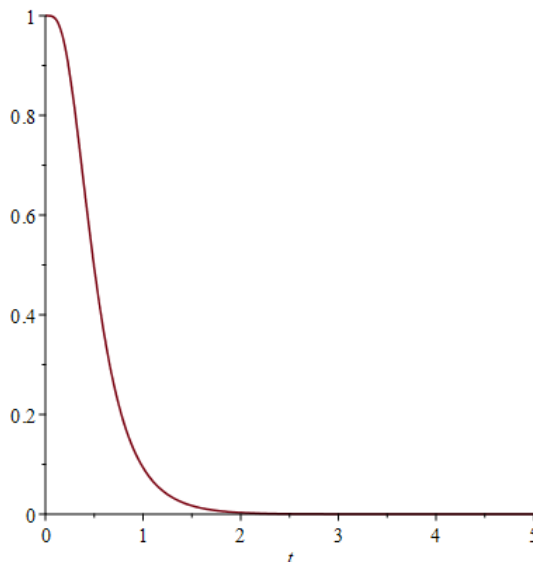


Fig. 4 The prediction of the system reliability function

Fig. 4 shows a graph of the prediction of the reliability function of a new system in the time interval  $tt = 2$  over time  $[0,5)$ .

The derivative  $dptt\_syst(t)$  of this function is found by the command:

$$dptt\_syst(t) := diff(ptt\_syst(t), t) :$$

Then we obtain the average operating time  $Ttt$  of the system, starting from the moment of time  $tt$ , over an interval of duration  $\tau$  by using the command:

$$Ttt := int(-t \cdot dptt\_syst(t), t = (0.) ..tau) + tau \cdot ptt\_syst(tau);$$



$$Ttt := 0.5653022542 .$$

The new system condition at a moment of time can be estimated by using the formula  $x := \frac{T0 - Ttt}{T0}$ . In this case  $x := 0.6066702896$ .

It is easy to obtain a prediction of the trajectory of the new system condition on the time interval  $[0, \tau)$ . To do this, we need to clear the variable  $tt$  by using the command  $tt := 'tt'$ . We define the function  $Ttt(tt)$  and function  $x(tt) = \frac{T0 - Ttt(tt)}{T0}$ . The graph of condition changes of the new system  $S$  on a time interval  $[0, \tau)$  can be obtained by using the command:

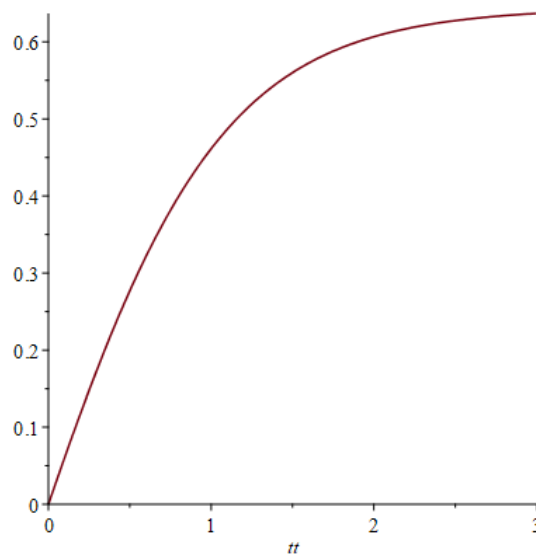
$$\text{plot}(x(tt), tt = 0 .. \tau);$$


Fig. 5 The dependence of the system condition on time

Fig. 5 shows a graph of the predicted condition changes of the new system  $S$  on the time interval  $[0, \tau)$ , where  $\tau = 3$ .

**3.** The set  $C$  of all subsets of a set of  $n = 9$  elements has cardinality  $2^n = 512$ . All subsets can be found in the MAPLE system by using the commands:

$$> \text{with}(\text{combinat});$$

$$> C := \text{choose}(9);$$

Let us group the resulting subsets into rectangular matrices by  $i$  elements in the row,  $i = 1, \dots, 9$ . We obtain matrices  $C_i, i = 1, \dots, 9$ , with a different number of rows. These matrices can be obtained sequentially by using the command:

$$> \text{convert}(\text{choose}(9, i), \text{Matrix});$$

The number of rows in the matrix  $C_i$  is equal to  $\frac{n!}{i!(n-i)!}$ . We assume that the  $k$ -th row of the

matrix  $C_i$  is a set of subsystem numbers. From the rows of the matrix  $C_i$ , we compose the matrix  $G_i$  according to the following rule. A row of the matrix  $C_i$  is included in the matrix  $G_i$ , if a failure of all subsystems listed in that row results in a failure of the system. The MAPLE program, which solves the problem of finding matrices for the considered system  $S$  is presented below. It has been found that in the model example there are 84 sets of subsystems, the failure of which leads to the system failure.

```

#Sets of subsystems, the failures of which leads to the system failure
>restart;
>with(combinat);

>n:=9; #A number of subsystems
>p:=convert(array(1..n),Vector);#Vector for storing the subsystem
#reliabilities
>L:=0; #A number of found sets
>for i from 1 to n
do
mi:=convert(choose(n,i),Matrix): #Matrix of sets of n elements by i
k_max:=n!/(i!*(n-i)!): #A number of all sets
for k from 1 to k_max
do
for l from 1 to n do p[l]:=1 od:
for I from 1 to I do p[mi[k,I]]:=0 od:
#System reliability calculation
p1:=(1-(1-p[2])*(1-p[4])*(1-p[6]))*(1-(1-p[3])*(1-p[5])*(1-p[7]))*p[8]*p[9]:
p2_s1:=(1-(1-p[2])*(1-p[4]))*(1-(1-p[3])*(1-p[5])):
p2_s2:=p[6]*p[7]:
p2:=(1-(1-p2_s1)*(1-p2_s2))*p[8]*(1-p[9]):
p3_s1:=p[2]*p[3]:
p3_s2:=(1-(1-p[4])*(1-p[6]))*(1-(1-p[5])*(1-p[7])):
p3:=(1-(1-p3_s1)*(1-p3_s2))*(1-p[8])*p[9]:
p4_s2:=p[6]*p[7]:
p4_s1:=1-(1-p[2]*p[3])*(1-p[4]*p[5]):
p4:=(1-(1-p4_s1)*(1-p4_s2))*(1-p[8])*(1-p[9]):
pp:=(p1+p2+p3+p4):
p_sys:=1-(1-p[1])*(1-pp):
If p_sys=0 then L:=L+1 fi:
If p_sys=0 then print(mi[k]) fi:
end do:
end do: print('A number of sets in G'L);

```

The basic set of failures  $M$  can be easily found by using the algorithm given in Section 2.5. The set  $M$  contains 8 sets of  $G_4$ ,  $G_5$ ,  $G_6$  :  $\{1,2,4,6\}$ ,  $\{1,3,5,7\}$ ,  $\{1,2,4,7,9\}$ ,  $\{1,2,5,7,8\}$ ,  $\{1,3,4,6,8\}$ ,  $\{1,3,5,6,9\}$ ,  $\{1,2,5,6,8,9\}$ ,  $\{1,3,4,7,8,9\}$ .

## REFERENCES

1. Barzilovich E.Yu., Belyaev Yu.K., Kashtanov V.A. and others. Ed. Gnedenko B.V. Mathematical theory of reliability issues. - M.: Radio and communication, 1983.[in Russian]
2. Kashtanov V.A., Medvedev A.I. "Theory of reliability of the difficult systems". M.: FIZMATLIT, 2010.[in Russian]
3. Herzbach I.B., Kordonsky H.B. Fault models. - M.: Sov. Radio, 1966. - 166p. [in Russian]
4. Barlow R.E., Proschan F. Mathematical Theory of Reliability. SIAM, 1996. — 258 p. [in Russian]
5. Tolok I.V. Mathematical model of technical maintenance of difficult refurbishable object without account of his structure. *Informatics and Mathematical Methods in Simulation*. Vol. 6, №4, P. 379-384. 2016 [in Ukrainian]
6. Podtseykin N.S. Mathematical model of the dynamics of restorable systems. *Radioelectronics & Informatics*". №4, P. 85-90. 2001.[in Russian]
7. Podtseykin N.S. A mathematical model for the prevention of complex technical systems. *Bulletin of V. Karazin Kharkiv National University, series Mathematical Modelling. Information Technology. Automated Control Systems*, Issue 31, P. 82-93, 2016. [in Russian]
8. Podtseykin N.S. Optimization of the reliability of a complex technical system. *Bulletin of V. Karazin Kharkiv National University, series Mathematical Modelling. Information Technology. Automated Control Systems*, Issue 39, P. 48-60. 2018.[in Russian]
9. H. Main, S. Osaki. Markov decision-making processes. - M.: Science, 1977. [in Russian]

10. Howard R.A. Dynamic programming and Markov processes. Sov. Radio, 1964. [in Russian]
11. A.L. Gorelik., V.A. Skripkin. Recognition methods, M.: Higher School, 2004. [in Russian]
12. Samarsky B.P., Gulin A.V. Numerical methods M: Science, 1989. – 432 p.
13. Rudin U. Functional analysis. M.: Mir, 1975. [in Russian]
14. Podtsykin N.S. Optimization of the observation period in the Markov decision-making process. Visnyk KhNU, №629, Seria “Mathematical Modeling. Information technology. Automated Control Systems”, Issue 3, Kharkiv, pp. 25-32, 2004. [in russian]
15. Podtsykin N.S. Optimization in periodic Markovian decision processes. Cybernetics and Systems Analysis, Kyiv, № 2, 1991, pp. 91-94,99. [in russian]
16. Dyakonov V.P. Maple 10/11/12/13/14 in mathematical calculations. - M.: DMK Press, 2011. 800 p. [in russian]

#### ЛІТЕРАТУРА

1. Барзилович Е.Ю., Беляев Ю.К., Каштанов В.А. и др. Под ред. Гнеденко Б.В. Вопросы математической теории надежности. – М.: Радио и связь, 1983. – 376с.
2. Каштанов В.А., Медведев А.И. Теория надежности сложных систем. – М.: ФИЗМАТЛИТ, 2010. – 608с.
3. Герцбах И. Б., Кордонский Х. Б. Модели отказов. - М.: Советское радио, 1966. - 166 с.
4. Барлоу Р., Прошан Ф. Математическая теория надежности. М.: Советское радио, 1969. - 488 с.
5. Толлок І.В. Математична модель технічного обслуговування складного відновлювального об'єкту без урахування його структури. *Інформатика та математичні методи в моделюванні*. – 2016. – Том 6, №4, с. 379-384.
6. Подцькин Н.С. Математическая модель динамики восстанавливаемых систем. Радиоэлектроника и информатика. 2001, №4, с. 85-90.
7. Подцькин Н.С. Математическая модель профилактики сложной технической системы. *Вісник Харківського національного університету ім. В.Н. Каразіна серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління»*. 2016. Вип. 31, С. 82-93.
8. Подцькин Н.С. Оптимизация надежности сложной системы. *Вісник Харківського національного університету ім. В.Н. Каразіна серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління»*, 2018. Вип. 39. С. 48-60.
9. Майн Х., Осаки С. Марковские процессы принятия решений. – М.: Наука, 1977. – 175с.
10. Ховард Р.А. Динамическое программирование и марковские процессы.: Сов. Радио, 1964. – 189 с.
11. Горелик А.Л., Скрипкин В.А. Методы распознавания, М.: Высшая школа, 2004. – 262с.
12. Самарский А.А, Гулин А.В. Численные методы. М: Наука, 1989. – 432 с.
13. Рудин У. Функциональный анализ. М.: Мир, 1975. – 443с.
14. Подцькин Н.С. Оптимизация периода наблюдений в марковском процессе принятия решений. *Вісник Харківського національного університету ім. В.Н. Каразіна серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління»*.2004. Вип. 3. С. 25-32.
15. Подцькин Н.С. Метод оптимизации в периодических марковских процессах принятия решений. Кибернетика, Киев, № 2, 1991, с. 91-94,99.
16. Дьяконов В.П. Maple 10/11/12/13/14 в математических расчетах. – М.: ДМК Пресс, 2011. – 800 с.

УДК 004.7(075)

## Модель процесу управління доступом у бездротовій комп'ютерній мережі

М.О. Слабишев, А.Г. Бердніков

**Слабишев  
Михайло Олегович**

студент кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022;

e-mail: [misha220498@gmail.com](mailto:misha220498@gmail.com);<https://orcid.org/0000-0003-2367-1731>**Бердніков  
Анатолій Георгійович**

доцент, кандидат технічних наук кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022;

e-mail: [a.berdnikov@karazin.ua](mailto:a.berdnikov@karazin.ua);<https://orcid.org/0000-0001-7923-0754>

В АСУ ТП до сьогодні існує проблема забезпечення мінімального часу доставки інформації при відсутності колізій, бо під час передачі часто виникають колізії або час затримки інформації є неоптимальним. Ця проблема вирішується використанням бездротових комп'ютерних мереж та методами передачі інформації з запобіганням колізії які впливають на час затримки передачі інформації. Розроблена модель процесу управління доступом в бездротових комп'ютерних мережах вирішує проблему ефективної передачі інформації з запобіганням колізії. Метою використання даної моделі є зменшення затримки передачі інформації.

*Ключові слова* : бездротова комп'ютерна мережа, колізія, кадр, слот, станція.

## Model of control access process in wireless computer network

M. Slabyshev, A. Berdnikov

**Slabyshev Mykhailo***student of the department of theoretical and applied systems engineering; V.N. Karazin Kharkiv National University, Maidan Svobody, 4, Kharkiv-22, Ukraine, 61022;***Berdnikov Anatoly***Associate Professor, Candidate of Technical Sciences, Department of Theoretical and Applied Systems Engineering; V.N. Karazin Kharkiv National University, Maidan Svobody, 4, Kharkiv-22, Ukraine, 61022;*

The problem of ensuring the minimum delivery time of information with prevention of collisions is a pressing problem in ACS TP, since collisions often occur during transmission or the delay time of information is not optimal. This problem is solved by the use of wireless computer networks and methods of information transmission with collision prevention, which affect the delay time of information transmission. The developed model of the access control process in wireless computer networks solves the problem of effective information transfer with collision prevention. The purpose of using this model is to reduce the delay in information transmission. ACS TP often needs modernization as technical innovations are becoming available. At the same time, the quality of production should not fluctuate. The most important element of modernization is the use of wireless technologies that save money and time, compared to the deployment of wired networks. The aim of the work is to develop the model of control access process in a wireless computer network with collision prevention. The process of forming delay slots for information transfer has been researched. The object of the study is a wireless computer network in the ACS TP. The subject of the study is a model of control access process in a wireless computer network with collision prevention. The research objectives are to create requirements for the model, to propose a structure, to create a model, to conduct tests and evaluate the obtained results. The model of access control process in a wireless computer network designed to transmit information with prevention of collisions in a wireless computer network has been developed.

*Keywords*: wireless computer network, collision, frame, slot, station, access control model, process, access mode.

## Модель процесса управления доступом в беспроводной компьютерной сети

М.О. Слабышев, А.Г. Бердников

**Слабышев  
Олегович****Михаил***студент кафедры теоретической и прикладной системотехники; Харьковский национальный университет имени В.Н. Каразина, площадь*

**Бердников  
Анатолий Георгиевич**

*Свободы, 4, Харьков-22, Украина, 61022;  
доцент, кандидат технических наук кафедры теоретической и прикладной  
системотехники; Харьковский национальный университет имени  
В.Н. Каразина, площадь Свободы, 4, Харьков-22, Украина, 61022;*

В АСУ ТП до сих существует проблема обеспечения минимального времени доставки информации при отсутствии коллизий, поскольку при передаче часто возникают коллизии или время задержки информации является неоптимальным. Эта проблема решается использованием беспроводных компьютерных сетей и методами передачи информации с предотвращением коллизии, которые влияют на время задержки передачи информации. Разработанная модель процесса управления доступом в беспроводных компьютерных сетях решает проблему эффективной передачи информации с предотвращением коллизии. Целью использования данной модели является уменьшение задержки передачи информации.

**Ключевые слова:** беспроводная компьютерная сеть, коллизия, кадр, слот, станция.

У статті розглянуті особливості процесу управління бездротових комп'ютерних мереж.

Метою даної роботи є розробка моделі процесу управління доступом у бездротовій комп'ютерній мережі з запобіганням колізій та методики вибору міжкадрових інтервалів. Обґрунтування вибору міжкадрових інтервалів в режимах доступу до середовища. Дослідження процесу формування слотів відстрочки для передачі інформації.

Об'єктом даної роботи є бездротова комп'ютерна мережа в АСУ ТП.

Предмет дослідження - модель процесу управління доступом у бездротовій комп'ютерній мережі з запобіганням колізій.

Розроблена модель процесу управління доступом у бездротовій комп'ютерній мережі, яка призначена для передачі інформації з запобіганням колізій в бездротовій комп'ютерній мережі.

Розроблена модель наочно показує роботу і особливості цього режиму. Необхідність розробки цієї моделі полягала в тому, що при вивченні режимів доступу в бездротовій мережі ця модель полегшила процес навчання і розуміння передачі кадрів в бездротових мережах.

Актуальність роботи: до сьогодні існує проблема передачі інформації в бездротових мережах, бо під час передачі часто виникають колізії. Тема є актуальною, бо розроблена модель протоколу управління доступом забезпечує запобігання колізії.

### **Аналіз проблем використання бездротових мереж**

АСУ ТП (Автоматизована система управління технологічним процесом) найчастіше потребує модернізації, так як технічний прогрес не стоїть на місці. При цьому, якість виробництва не повинна змінюватися і має бути на однаково високому рівні. Найважливішим елементом модернізації є застосування бездротових технологій, що приносять економію коштів і часу, в порівнянні з розгортанням дротових мереж. [1]

У більшості застосувань бездротові мережі дозволяють досягти наступних переваг в порівнянні з провідними мережами:

- суттєво знизити вартість встановлення датчиків;
- зменшити затримку передачі інформації;
- зменшити трудовитрати часу на монтаж і обслуговування системи;
- прискорити налагодження системи і пошук несправностей;
- забезпечити зручну модернізацію системи.

Таким чином, застосування бездротових мереж є в деяких випадках більш зручним, ніж використання дротових технологій.

Характерною особливістю бездротових комп'ютерних мереж є те, що вони використовують не множинний доступ з виявленням колізій (CSMA / CD - Carrier Sense Multiple Access with Collision Detection.), а множинний доступ з запобіганням колізій (CSMA / CA - Carrier sensing multiple access with collision avoidance.). [2]

Стандарт бездротових мереж 802.11 намагається уникати колізій за рахунок множинного доступу з запобіганням колізій CSMA/CA. Запобігання колізій використовується для того, щоб поліпшити продуктивність передачі, віддавши мережу єдиному передавальному пристрою. Збільшення ефективності досягається за рахунок зниження ймовірності колізій і повторних спроб передачі. Уникнення колізій корисно на практиці в тих ситуаціях, коли своєчасне виявлення колізії неможливо - наприклад, при використанні радіопередавачів. [3]

В даному методі розрізняють 2 режими доступу до середовища: розподілений (DCF - **Distributed Coordination Function**) і централізований (PCF - Point Coordination Function.).

#### Розподілений режим доступу DCF

В цьому режимі реалізується безпосередньо метод CSMA / CA. Кожен переданий кадр повинен бути підтверджений кадром позитивної квитанції, який надсилається станцією призначення. Якщо по закінченню тайм-ауту для цього кадру квитанція не надходить, станція-відправник вважає, що сталася колізія.

Режим доступу DCF потребує синхронізації станцій. У стандарті 802.11 ця проблема вирішується так - тимчасові інтервали починають відраховуватися від моменту закінчення передачі попереднього кадру.

Станція, яка хоче передати кадр, зобов'язана попередньо перевірити зайнятість середовища. Як тільки вона підтверджує закінчення передачі попереднього кадру, вона зобов'язана відрахувати інтервал часу, рівний міжкадровому інтервалу IFS (Inter Frame Space). [4]

Якщо після закінчення IFS середовище все ще вільне, то починається відлік слотів, визначених заздалегідь тривалості (як зазначено на рисунку 1). Кадр можна починати передавати тільки на початку слота за умови, що середовище передачі в цей самий момент вільне.

Розмір слота залежить від способу кодування сигналу: для методу DSSS, який використовується саме в бездротових мережах – 1 мкс. Розмір слота вибирається так, щоб він перевершував час поширення сигналу між будь-якими двома станціями мережі плюс час, що витрачається станцією на розпізнавання ситуації зайнятості середовища. Якщо це дотримується, то кожна станція мережі зможе правильно розпізнати початок передачі кадру при прослуховуванні середовища. Це, в свою чергу, означає, що колізія може статися тільки в тому випадку, коли кілька станцій вибирають один і той самий слот для передачі.

Станція вибирає для передачі слот на підставі усіченого експоненціального двійкового алгоритму відстрочки (алгоритм додавання перевірочних слотів перед передачею основного кадру). Номер слота вибирається як випадкове ціле число, яке рівномірно розподілене в інтервалі  $[0, CW]$ , де  $CW$  означає **Contention Window** (конкурентне вікно, що може змінюватися в розмірах – 2,4,8,16 або 32). [5]

Для наочності прикладу роботи режиму DCF розроблена модель, яка наочно показує роботу і особливості цього режиму. Необхідність розробки цієї моделі полягала в тому, що при вивченні режимів доступу в бездротовій мережі ця модель полегшує процес навчання і розуміння передачі кадрів в бездротових мережах. Дана модель може бути використана для мережі з великою кількістю станцій (до 32 станцій), в залежності від розміру конкурентного вікна.

Припустимо ситуацію, коли станція В хоче передати кадри станції А. Перед передачею кадрів і відрахування слотів відстрочки, станція, яка хоче захопити середовище, обов'язково повинна послати службовий кадр RTS (Request To Send- запит на передачу). На цей кадр станція, яка буде приймати кадри, повинна у відповідь послати службовий кадр CTS (Clear To Send - вільна для передачі). Тільки після цього починається алгоритм відліку слотів.

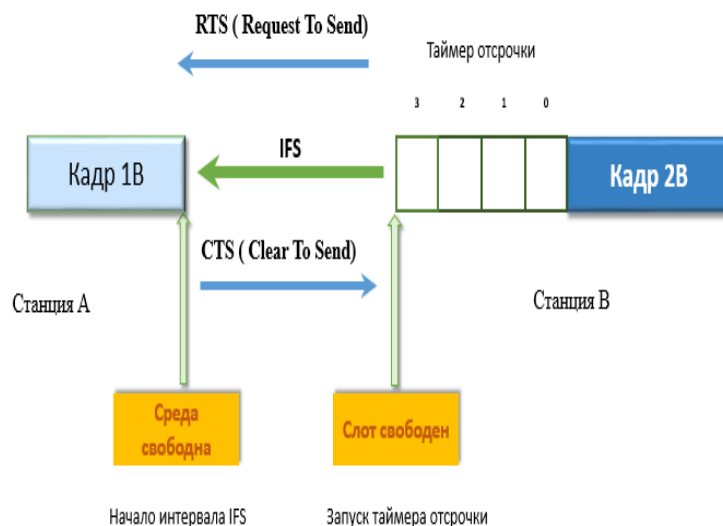


Рисунок 1 – Модель після першого кроку алгоритму

Перед передачею кадру станція В відлічить слоти відстрочки, починаючи зі слота 3. Після успішного захоплення середовища починається алгоритм зменшення слотів.

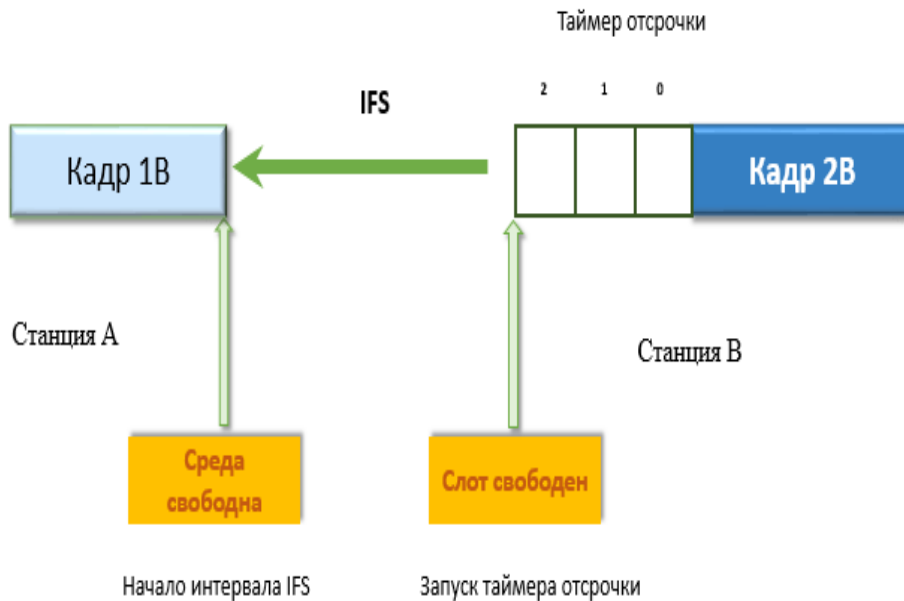


Рисунок 2 – Модель після другого кроку алгоритму

Так як середовище є вільним, слот є незайнятим, то кількість слотів успішно зменшується на 1.

Припустимо, в цей же час приєднується станція С, яка хоче передати кадри на станцію А. Станція С відрахувала для себе випадкове число слотів відстрочки  $i$ , припустимо, воно дорівнює 2 слотів (починаючи зі слота №1). Вона повинна послати службовий кадр RTS на станцію А і дочекатися відповідного службового кадру CTS.

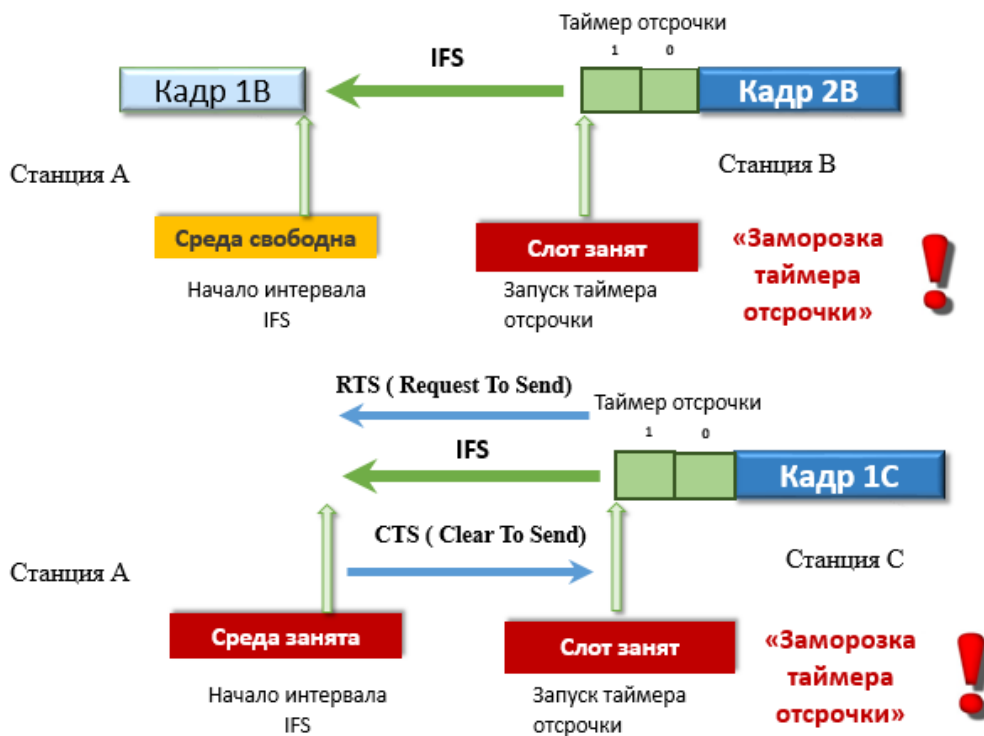


Рисунок 3 – Модель після третього кроку алгоритму

На малюнку зображена ситуація, коли станція В успішно відраховує слоти відстрочки для передачі кадру, перебуваючи на слоті № 1. Станція С робить запит на станцію А і отримує відповідь, що навколишнє середовище є зайнятим.

В даному випадку відбувається заморозка таймера відстрочки і для станції В, і для станції С, так як вони знаходяться на однаковому слоті відстрочки (слоті №1).

Так як для станції С середовище є зайнятим, їй необхідно відррахувати нову кількість слотів за формулою(1):

$$CW_{new} = \min \{2 (CW_{old} + 1) - 1, CW_{max}\}, \quad (1)$$

де  $CW_{max}$  - максимальний розмір конкурентного вікна;

$CW_{old}$  - попередня кількість слотів;

$CW_{new}$  - нова кількість слотів.

Лічильник для станції В просто заморожується, так як для неї середовище є вільною, а зайнятим є лише слот відстрочки.

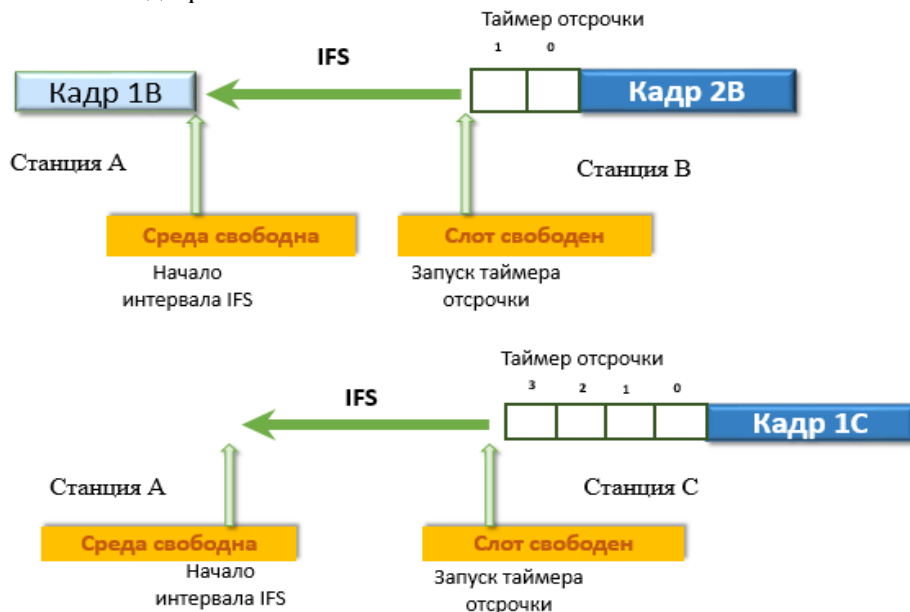


Рисунок 4 – Модель після четвертого кроку алгоритму

Для станції С порахували нову кількість слотів відстрочки, відмінне від кількості слотів на станції В. Для неї тепер є вільними і середовище передачі, і сам слот.

На станції В лічильник слотів відстрочки розморожується, так як її кількість слотів не збігається з кількістю слотів на іншій станції.

Станції готові для подальшої операції передачі кадру, попередньо відрахувавши свої слоти відстрочки.

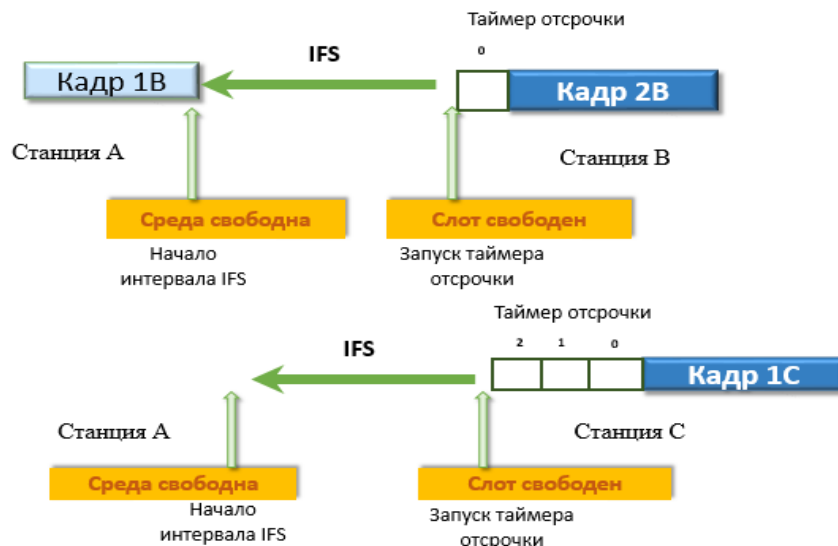


Рисунок 5 – Модель після п'ятого кроку алгоритму



Станція В відрахувала свій слот відстрочки для станції А, а станція С відрахувала свій.

Алгоритм буде виконуватися до тих пір, поки станція успішно не передасть свої кадри на приймальну станцію.

Дана імітаційна модель застосовує розподілений режим доступу до середовища при передачі інформації.

У даній моделі підраховується час очікування передачі кадрів. Він обчислюється за формулою (2):

$$t_{ож} = 2 * \frac{D}{V_c} + t_3 + t_{a1} + t_{a2} \quad , \quad (2)$$

де  $D$  - це відстань між приймаючої і відправна станціями;

$V_c$ - швидкість поширення сигналу (є сталим виразом);

$t_3$ - тривалість сигналу запиту (дорівнює розміру одного слота - 1 мкс);

$t_{a1}$  і  $t_{a2}$ - час обробки сигналу на приймаючій і передавальній станції.

Так як  $t_{a1}$  і  $t_{a2}$  вимірюються в наносекундах, при підрахунку цими значеннями можна нехтувати (при роботі в моделі ці обчислення не враховуються). Необхідно тільки знати відстань між станціями.

Отже, цей метод успішно працює і забезпечує запобігання колізій під час передачі кадрів між станціями. Під час роботи моделі можна розрахувати необхідну кількість слотів відстрочки для станції-передавача та час очікування передачі інформації. Модель запобігає колізії, що пришвидшує час передачі інформації від передавача до приймача.

Завдяки формулі (3) знаходиться мінімальна кількість кодових комбінацій, яку необхідно зберегти в пам'ятовуючому приладі(ЗП)(об'єм ЗП):

$$M = 1 + \left( \frac{t_{ож}}{(m+k)*T_c} \right), \quad (3)$$

де  $m + k$  – число інформаційних та перевірочних розрядів в кодовій комбінації;

$T_c$ - час обробки сигналу;

Час обробки сигналу з формули (3) не є сталою величиною та підраховується за формулою(4):

$$T_c = \frac{1}{B} - \text{час обробки сигналу}; \quad (4)$$

де  $B$  – швидкість модуляції (дорівнює 2400 біт/с).

Рахується також максимальна відстань, за якою є можливим передача даних між станціями:

$$D_{max} = (M - 1) * \frac{(m+k)*V_c}{2*B} - \frac{t_3 + t_{a1} + t_{a2}}{2} * V_c. \quad (5)$$

Якщо  $D_{max} < D$ , то можна вважати, що передача інформації неможлива, бо відстань для передачі є більшою, ніж відстань, на яку можлива ця передача. [6]

Так як  $t_{a1}$  і  $t_{a2}$  вимірюються в наносекундах, при підрахунку цими значеннями можна нехтувати (при роботі в моделі ці обчислення не враховуються). Необхідно тільки знати відстань між станціями.

### 3. Результати дослідження роботи моделі

Дослідження моделі полягло в тому, щоб розрахувати час очікування інформації при різних відстанях між станцією-передавачем і приймаючою станцією. Завдяки отриманому часу очікування знаходиться максимальна відстань між станціями. Цей результат дає зрозуміти доцільність використання саме бездротових технологій в АСУ ТП.



Рисунок 6 – Залежність часу очікування передачі від максимальної відстані передачі інформації

Аналіз дослідження моделі показав, що час очікування передачі інформації залежить від максимально можливої відстані передачі інформації. Чим більша буде максимально можлива відстань передачі, тим більший час буде затрачений на передачу інформації. Є доцільним використання саме бездротових технологій в АСУ ТП, бо саме такий метод передачі інформації дозволяє передавати інформацію на великі відстані, не застосовуючи дротових технологій, а час, витрачений на передачу інформації, є оптимальним.

### Висновки

Отже, є доцільним використання в АСУ ТП саме бездротових технологій, так як вони є більш економічно вигідними, ніж дротові технології. Метод передачі інформації з запобіганням колізій забезпечують безпомилкову передачу даних на відносно далекі відстані за оптимальний час передачі.

Необхідність розробки такої моделі полягала у використанні методів передачі інформації з запобіганням колізій. Саме цей метод вважається найефективнішим, бо він потребує меншого часу, аніж в інших методах. А час, витрачений на передачу інформації, є критично важливою характеристикою при виборі технологій та методів передачі даних.

Дана модель дозволяє використовувати бездротову мережу для великої кількості станцій (до 32 станцій), яка залежить від розміру конкурентного вікна, розміри якого можуть змінюватись. Чим більший розмір конкурентного вікна, тим більше станцій зможуть обмінюватись інформацією в однаковий проміжок часу.

Таким чином, розроблена методика вибору міжкадрових інтервалів, що дозволяє запобігти виникненню колізій, що, в свою чергу, впливає на час затримки та передачі інформації та несе за собою вплив на цінність інформації.

## ЛІТЕРАТУРА

1. Стаття про приклад використання бездротових технологій в промисловості: веб-сайт.  
URL:<http://1234g.ru/wifi/standarty-wifi> (Дата звернення: 02.07.2020)
2. Ю. І. Лосев, Руккас К. М., Шматков С. І. Комп'ютерні мережі : навчальний посібник / ХНУ імені В. Н. Каразіна, 2013. – 248 с.  
<https://www.univer.kharkov.ua/images/redactor/news/2014-04-16/Losev.pdf>
3. Блок-схема алгоритму CSMA/CA при передаванні даних: веб-сайт.  
URL:[https://neerc.ifmo.ru/wiki/index.php?title=Data link layer - MAC - Aloha, CSMA](https://neerc.ifmo.ru/wiki/index.php?title=Data_link_layer_-_MAC_-_Aloha,_CSMA)  
(Дата звернення: 02.06.2020)
4. Оліфер В., Оліфер Н. Комп'ютерні мережі. Принципи, технології, протоколи : Підручник для внз. 4-е видавництво – СПб.: Пітер, 2016. – 992 с.: мул. – (Серія «Підручник для внз»).  
<http://naymov.com/edu/ukit/olifer.pdf>
5. Стаття про підрахунок слотів в конкурентному вікні: веб-сайт.  
URL:<https://mipt.ru/drec/upload/d27/multihop-wireless.pdf> (Дата звернення: 28.05.2020)
6. Лосев Ю.І., Плотников Н. Д. Методи передачі інформації.: навчальний посібник / Вірта ім. Л.А.Говорова, 1978р. – 187 с.

## REFERENCES

1. An article on an example of the use of wireless technologies in industry[Online].  
Available: <http://1234g.ru/wifi/standarty-wifi>  
Accessed on : 02.07.2020 [in Russian]
2. U.I. Losev, K.M. Rukkas, S.I. Shmatkov, *Computer networks: a textbook* / VN Karazin KhNU, 2013. - 248 p.  
<https://www.univer.kharkov.ua/images/redactor/news/2014-04-16/Losev.pdf> [in Ukrainian]
3. Block diagram of the CSMA / CA algorithm for data transmission[Online].  
Available: [https://neerc.ifmo.ru/wiki/index.php?Title=Data link layer - MAC - Aloha CSMA](https://neerc.ifmo.ru/wiki/index.php?Title=Data_link_layer_-_MAC_-_Aloha_CSMA)  
Accessed on: 02.06.2020 [in Russian]
4. V.Olifer, N.Olifer, *Computer networks. Principles, technologies, protocols: A textbook for universities*. 4th publishing house - St. Petersburg: Peter, 2016. - 992 p .: mul. - (Series "Textbook for universities")  
<http://naymov.com/edu/ukit/olifer.pdf> [in Russian]
5. An article about counting slots in a competitive window[Online].  
Available: <https://mipt.ru/drec/upload/d27/multihop-wireless.pdf>  
Accessed on: 28.05.2020 [in Russian]
6. Y.I. Losev, N.D. Plotnikov. *Methods of information transfer .: textbook* / : Virta L.A. Govorov, 1978 - 187 p.[in Russian]

УДК 004

## The model of accounting for the research activity of education system employees based on the IOS operating system

Olena Tolstoluzka, Krystyna Afanasieva

**Tolstoluzka Olena**

*professor of the Department of Theoretical and Applied Systems Engineering  
Kharkiv National University named after V. N. Karazin, Kharkiv,  
Svobody Square, 4, 61022*

*e-mail: elena.tolstoluzka@karazin.ua;*

<https://orcid.org/0000-0003-3085-2346>

**Afanasieva Krystyna**

*student of the Department of Theoretical and Applied Systems Engineering  
Kharkiv National University named after V. N. Karazin, Kharkiv, Svobody  
Square, 4, 61022*

*e-mail: kristina.afanasieva.dev@gmail.com*

<https://orcid.org/0000-0002-6262-4818>

The process of reviewing and analyzing of scientific information resources has been discussed in the paper. Analysis of software solutions, as well as, the conclusions as for the possibility of developing additional information arrays based on existing sources of scientific and scientometric information have been presented. The paper offers the approaches to storing scientometric information, the optimized methods of obtaining and processing data. That provides the ability to digitize publications to simplify searching and filtering scientific information and its further processing and replenishment of information arrays with a new data. The architectural templates have been analyzed and the Viper has been selected as the most suitable architectural template for the distribution of responsibilities and test coverage of the application. The choice of Swift programming language has been substantiated. Strengths and weaknesses of these technologies have been considered. The OS and Firebase data storage tool have been selected and discussed. The advantages of those technologies for the development of a mobile application have been listed. The purpose of the work is to create a mobile application for the operating system "iOS", version 11.0. The application must be a client that works with Firebase Cloud Firestore, which allows searching and filtering publications, as well as, adding and deleting records from Firebase Cloud Firestore for the administrators. Therefore it is necessary to implement the functionality that allows registering new and authorizing registered users. The object of the research is processing scientometric and scientific knowledge. The subject of research is the methods and tools of technology to automate the process of knowledge recording based on the scientific and scientometric information.

**Keywords:** filter, search, optimization, subject area, Scopus, Web of Science, Swift, iOS, application, Firebase, CocoaPods, operating system.

## Модель урахування дослідницької діяльності працівників системи освіти на базі операційної системи IOS

О. Г. Толстолузка, Х. О. Афанасьєва

**Толстолузка**

**Олена Геннадіївна**

*д.т.н., с.н.с., професор кафедри теоретичної і прикладної системотехніки  
факультету комп'ютерних наук;*

*Харківський національний університет імені В. Н. Каразіна,*

*майдан Свободи, 4, Харків, Україна, 61022*

*e-mail: elena.tolstoluzka@karazin.ua;*

<https://orcid.org/0000-0003-3085-2346>

**Афанасьєва**

**Христина Олександрівна**

*студентка кафедри теоретичної та прикладної системотехніки;*

*Харківський національний університет імені В. Н. Каразіна, м. Харків,  
майдан Свободи, 4, 61022*

*e-mail: kristina.afanasieva.dev@gmail.com*

<https://orcid.org/0000-0002-6262-4818>

Робота присвячена огляду та аналізу ресурсів наукової інформації, аналізу засобів програмних рішень для досягнення поставленої мети, дослідження можливостей розробки додаткових інформаційних масивів на базі існуючих джерел наукової та наукометричної інформації. В рамках роботи запропоновані підходи зберігання наукометричної інформації, способи отримання та обробки даних, що дозволить оптимізувати час роботи з інформацією та надасть можливість оцифровки публікацій, що дозволяє залучати менше ресурсів для реалізації робіт з пошуку та фільтрації наукової інформації і подальшої її обробки та поповнення інформаційних масивів новими даними. В роботі були

проаналізовані архітектурні шаблони після чого зроблено висновок, що для найкращого розподілу обов'язків, і покриття додатку тестами, найбільше підходить архітектурний шаблон - VIPER. Обґрунтований вибір мови програмування Swift. Було розглянуто сильні та слабкі сторони при використанні цих технологій. Розглянуто обрану ОС та засіб збереження даних Firebase. Перераховані переваги вищеперахованих технологій для розробки мобільного додатку. Метою роботи є створення мобільного додатку для операційної системи "iOS" компанії Apple, версії вище 11.0. Додаток має являти собою клієнт який працює з Firebase Cloud Firestore, який дозволяє здійснювати пошук та фільтрацію публікацій а також у випадку, якщо користувач є адміністратором додавати та видаляти записи з Firebase Cloud Firestore. В додатку необхідно реалізувати функціонал, що дозволяє виробляти реєстрацію і авторизацію зареєстрованим користувачам. Об'єктом дослідження є процес обробки наукометричних та наукових знань. Предметом дослідження є методи і засоби інформаційних технологій для автоматизації процесу обліку знань на базі ресурсів наукової та наукометричної інформації.

**Ключові слова:** *фільтр, пошук, оптимізація, предметна область, Scopus, Web of Science, Swift, iOS, application, Firebase, CocoaPods, операційна система.*

## Модель учета исследовательской деятельности работников системы образования на базе операционной системы IOS

Е. Г. Толстолужская, К. А. Афанасьева

Толстолужская  
Елена Геннадьевна

*д. т. н., с. н. с. профессор кафедры теоретической и прикладной системотехники факультета компьютерных наук  
Харьковский национальный университет имени В. Н. Каразина  
г. Харьков, площадь Свободы, 4, 61022  
e-mail: elena.tolstoluzka@karazin.ua;*

<https://orcid.org/0000-0003-3085-2346>

Афанасьева  
Кристина  
Александровна

*студентка кафедры теоретической и прикладной системотехники  
Харьковский национальный университет имени В. Н. Каразина  
г. Харьков, площадь Свободы, 4, 61022  
e-mail: kristina.afanasieva.dev@gmail.com*

<https://orcid.org/0000-0002-6262-4818>

Работа посвящена обзору и анализу ресурсов научной информации, анализа средств программных решений для достижения поставленной цели, исследования возможностей разработки дополнительных информационных массивов на базе существующих источников научной и наукометрической информации. В рамках работы предложены подходы хранения наукометрической информации, способы получения и обработки данных, что позволит оптимизировать время работы с информацией и позволит оцифровки публикаций, позволяет привлекать меньше ресурсов для реализации работ по поиску и фильтрации научной информации и последующей ее обработки и пополнение информационных массивов новыми данными. В работе были проанализированы архитектурные шаблоны после чего сделан вывод, что для лучшего распределения обязанностей, и покрытие приложении тестами, больше подходит архитектурный шаблон - VIPER. Обоснован выбор языка программирования Swift. Были рассмотрены сильные и слабые стороны при использовании этих технологий. Рассмотрены выбранную ОС и средство хранения данных Firebase. Перечисленные преимущества вышперечисленных технологий для разработки мобильного приложения. Целью работы является создание мобильного приложения для операционной системы "iOS" компании Apple, версии выше 11.0. Приложение должно представлять собой клиент работающий с Firebase Cloud Firestore, который позволяет осуществлять поиск и фильтрацию публикаций а также в случае, если пользователь является администратором добавлять и удалять записи из Firebase Cloud Firestore. В приложении необходимо реализовать функционал, позволяющий производить регистрацию и авторизацию зарегистрированным пользователям. Объектом исследования является процесс обработки наукометрических и научных знаний. Предметом исследования являются методы и средства информационных технологий для автоматизации процесса учета знаний на базе ресурсов научной и наукометрической информации.

**Ключевые слова:** *фильтр, поиск, оптимізація, предметная область, Scopus, Web of Science, Swift, iOS, application, Firebase, CocoaPods, операционная система.*

### 1. Introduction

The rapid development of information technology, its introduction into all spheres of modern life has led to a sharp expansion of the concept of knowledge processing in computer systems. At the beginning of the last century the total amount of human knowledge doubled about every fifty years, nowadays it doubles every five years. That is why the problem of creating methods that will help rationally organize the process of storing and operating information is of great importance.

Every year, the volume of scientific information grows. The number of experiments and publications in various subject areas has been increasing constantly, resulting in an ever-growing need for recording of scientometric information.

The development and use of scientometric knowledge by modern digital technologies will form a comprehensive view of the analysis of publishing activity and the possibility of using obtained data to make important management decisions.

It is difficult to imagine a modern man without a mobile device in his hands. Whether it's a phone, smartphone or other communicators - these devices are firmly and, apparently, for a long time, have become established in society as indispensable everyday means of communication. With the advent of new mobile devices and the variety of platforms on which they are developed, the business of creating mobile applications for various purposes is expanding actively.

Due to the spread of COVID-19 we are facing with the issue of distancing all educational and scientific processes, and recording the scientific information is not an exception.

## 2. Formulation of the problem

The purpose of the work is to create a mobile application for optimizing information about the scientific activities of the education system employees based on the IOS operating system.

To achieve this goal, the following tasks have been formulated:

- to analyze the literature on existing software solutions and choose technologies for creating a mobile application;
- to create a data storage model using Firebase;
- to develop a mobile application for optimizing the research activities of the education system employees by using the Swift programming language;
- to test the developed software;
- to create the documentation for the developed application;
- to analyze the quality and speed of test results of the developed application;
- to form an explanatory note.

## 3. Main part

The application prototype is the computer model of accounting for the scientific activities of the education system employees created by using the C++ language, which has been developed as a bachelor's thesis.

The main advantages of mobile applications are the simplicity of the communication between the brand and the user, economic benefits and ease of use.

Interactive versions of printed publications are becoming more and more popular, which is why we would like to overcome the limitations of working only with publications that are part of Scopus or Web of Science and cover most of this subject area.

A mobile application can be an effective tool for attracting new students to participation in scientific conferences or a convenient for working with existing publications.

### **The developed product has the following functionality:**

- Integration with Firebase, namely:
  - Google Authorization.
  - Authorization via Facebook.
  - Authorization via mail.
  - Data storage.
- The authorization not only for administrators but also for regular users, which will allow user-friendly functionality, has been added.
- The automatic check for belonging to the administrators' list has been implemented for authorization via mail. Successful check provides access to the functionality intended for the administrators only, namely:
  - Adding new records with the possibility to attach a document.
  - Deleting old records.
- Adding publications to the favorites list.

- Viewing all posts.
- Filtering the mail.
- Searching for the publications by specified parameters.
- Downloading the document.

#### 4. Choice of programming language

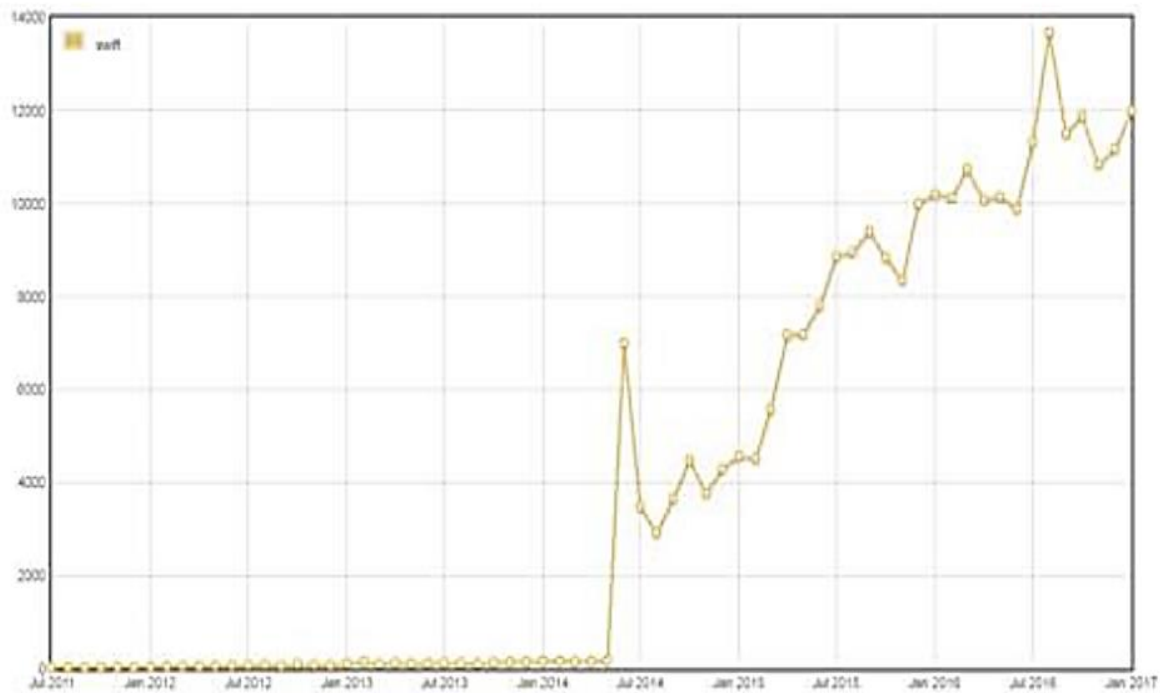
After analyzing the advantages and disadvantages of Swift and Objective-C, Swift has been selected as more convenient. Swift is a more productive language that allows faster creating of new applications. The main differences between these languages are presented in Table 1.

Table 1 - Comparison of programming languages Swift and Objective-C

	Swift	Objective-C
Accessibility to users of other programming languages	The syntax is similar to other well-known programming languages such as: Python, Go, JavaScript.	The syntax is much different from other programming languages, so it is quite difficult to learn.
Code length	The code is more streamlined, which reduces its length and development time.	The implementation of the same functions requires 2 times more lines of code than in Swift.
Interactivity	You can create applications interactively.	You cannot create applications interactively.
Community using the language	Swift is much easier to learn, but there are far fewer professionals now than in Objective-C.	Most of the iOS applications in the industry are written in Objective-C.
Safety	Security is one of the core concepts of Swift. Its design eliminates several types of potential errors in Objective-C.	Not as secure as Swift.
Open source	Swift is an open source language. It is also used in Linux, which creates the possibility of its application in the development of programs for Android OS.	Has no source code, and is exclusively focused on Apple devices.
Completeness and stability of the language	Swift is still under development and Apple is constantly releasing and adding new features.	Objective-C has already been created and used consistently.

When using Swift, the development process is simplified, easily accessible for novice programmers, thereby reducing the cost of highly qualified specialists. The Apple support for language development plays an important role as well. Rapid growth of the Swift community and its open-source character

should be mentioned. Swift has been chosen because it is free and therefore available to a large number of developers. Fig. 1 shows a graph of the growth of the Swift developer community since its launch.



*Fig. 1 - Swift community growth graph*

Reducing the time to write an application allows you to reduce costs and increase the company's profit. Swift is suitable for writing new applications in short time.

## **5. Choosing an Application Architecture Pattern**

As a result of the analysis of architectural patterns, the following conclusions can be made:

- 1) Traditional MVC is not applicable to modern IOS development.
- 2) Apple MVC is the best architectural pattern, in terms of development speed, but noticeably complicates the application support.
- 3) MVP offers good testability, but also means a significant increase in code size. Its features lie in the distribution because most of the responsibilities are divided between the Presenter and the Model, while the View has a minimum of tasks. Testability is also an advantage that is you can test most of the business logic because of the practical empty View. MVP, as well as traditional MVC, is not suitable for the IOS development.
- 4) The peculiarities of MVVM are in propagation. The View in MVVM has more responsibilities than the View in MVP, since MVVM updates the state from the ViewModel, and the MVP only passes all events to the Presenter and does not update itself. In testability the ViewModel knows nothing about the View, which allows you to test it without problems. The View can also be tested, but since it depends on UIKit it is not advisable to test it. Thus, MVVM combines the advantages of all the previously considered approaches to the architecture of IOS applications and it does not require additional code to update the View. Testability is also advantageous.



5) VIPER surpasses all presented architectural patterns in distribution, testability, and ease of learning new modules. The main features of VIPER are distribution, namely an excellent ability to assign responsibilities, and testability. It is ease of use as well. A small amount of code in VIPER blocks allows easily switching from one task to another, thereby increasing the usefulness of the programmers' work.

We can conclude that for the best distribution of responsibilities, and tests coverage of the application, the VIPER architectural pattern is the most suitable.

## 6. Conclusions

The main functions and capabilities of the Firebase database have been analyzed, and the Swift programming language has been chosen for developing a mobile application by using the VIPER architectural pattern.

## REFERENCES/ЛІТЕРАТУРА

1. Swift. Open Source Programming Language [Electronic resource]. - access mode: URL: <https://www.apple.com/ru/swift/> (date of access - 04.10.2020)
2. Introduction to Obj-C [Electronic resource]. - access mode: URL: <http://macbug.ru/cocoa/objc> (date of access - 17.10.2020)
3. React Native Tutorial: Build app with JavaScript [Electronic resource]. - access mode: URL: <https://www.raywenderlich.com/126063/reactnative-tutorial> (date of access - 15.10.2020)
4. The MVC design pattern [Electronic resource]. - mode Access: URL: <https://webformyself.com/shablon-proektirovaniya-mvc/> (date requests - 17.10.2020)
5. Model-View-Controller (MVC) in IOS: A Modern Approach [Electronic resource]. - access mode: URL: <https://www.raywenderlich.com/132662/mvc-in-ios-a-modern-approach> (date requests - 17.10.2020)
6. IOS Architecture Patterns [Electronic resource]. - access mode: URL: <https://techblog.badoo.com/blog/2016/03/21/ios-architecture-patterns/> (date requests - 17.10.2020)
7. Implementation of MVVM in IOS using RxSwift [Electronic resource]. - access mode: URL: <https://habrahabr.ru/post/273455/> (date requests - 17.10.2020)
8. The Book of VIPER [Electronic resource]. - access mode: URL: <https://github.com/strongself/The-Book-of-VIPER> (date accessed - 17.10.2020)
9. ANON The Swift Programming Language (Swift 2.1) / ANON – Cupertino: Apple Inc., 2014. 528 p.
10. Vandad Nahavandipoor iOS 8 Swift Programming Cookbook / Vandad Nahavandipoor – Boston: O'Reilly Media., 2014. – 902 p.

УДК 681.5.01

## Model of process control system in greenhouse agro-industrial complex

Oleksandr Sharapa, Anatoly Berdnikov

**Oleksandr Sharapa**

*student, 6th year student of the Department of Theoretical and Applied Systems Engineering, V.N. Karazin Kharkiv National University, 4 Svobody Square, Kharkiv, Ukraine, 61022;*

**Anatoly Berdnikov**

*Cand. of techn. sc., Associate Professor; Associate Professor of the Department of Theoretical and Applied Systems Engineering, V.N. Karazin Kharkiv National University, 4 Svobody Square, Kharkiv, Ukraine, 61022.*

The model of an automated control system for microclimatic indicators and energy consumption of greenhouses has been developed. To design intelligent components and on their basis to synthesize adaptive computer systems for managing the microclimate and energy efficiency of greenhouses is an urgent problem. The purpose of the work is to analyze greenhouse management modes, to make comparative characteristic of the process automation systems, to develop an information scheme of greenhouse-environment interaction, as well as, a structural scheme of greenhouse microclimate control, to select modern technologies to implement automatic climate control systems and to analyze the capabilities of the developed control model of a technological process. The research methods are based on the basic principles of the theory of heat and mass transfer using modern elements of computerized control. The structure of the greenhouse microclimate control system with a variable composition of equipment that provides rapid adaptation to the management requirements of a particular greenhouse has been developed. It has been proposed to develop the components and the structure of a greenhouse microclimate control system with implementation of the comprehensive approach, which includes communication and information management systems and technologies, a modern element base, the Android software, decision support tools. This approach is based on the following principles: consistency, variability of the equipment composition, openness, modularity, and usage of a set of basic design solutions. The microcontroller is the central control device of the entire control system. An availability graph of the client-server communication for the greenhouse microclimate control system, which has two final states, has been developed.

**Keywords:** *microclimate, greenhouse, control, automatic control system, microprocessor, hardware, element base.*

## Модель системи управління технологічним процесом в тепличному агропромисловому комплексі

О.В. Шарапа, А.Г. Бердников

**Шарапа Олександр  
Вадимович**

*студент, студент 6-го курсу кафедри теоретичної і прикладної системотехніки, Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків, Україна, 61022;  
e-mail: [sasha10ok64@gmail.com](mailto:sasha10ok64@gmail.com)  
<https://orcid.org/0000-0003-0284-2802>*

**Бердников Анатолій  
Георгійович**

*к. т. н., доцент; доцент кафедри теоретичної і прикладної системотехніки, Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків, Україна, 61022;  
e-mail: [a.berdnikov@karazin.ua](mailto:a.berdnikov@karazin.ua)  
<https://orcid.org/0000-0001-7923-0754>*

Розробка моделі автоматизованої системи управління мікрокліматичними показниками і енергоспоживанням тепличного господарства. Актуальною проблемою є проектування інтелектуальних компонентів та синтез на їх основі адаптивних комп'ютерних систем управління мікрокліматом та енергоефективністю тепличного господарства. Метою роботи є аналіз методів управління режимами роботи тепличного господарства, порівняльна характеристика засобів автоматизації управління технологічним процесом, розробка схеми інформаційної взаємодії теплиці з оточенням та структурної схеми управління мікрокліматом теплиці, підбір сучасних технічних засобів реалізації систем автоматичного регулювання мікроклімату та дослідження можливостей розробленої моделі управління технологічним процесом. Методи досліджень ґрунтуються на основних положеннях теорії тепло- і масообміну з використанням сучасних елементів комп'ютеризованого управління. Розроблено структуру системи управління мікрокліматом у теплиці, яка має змінний склад обладнання, що забезпечує швидку адаптацію до вимог управління для конкретної теплиці. Запропоновано розробляти компоненти і синтезувати структуру системи управління мікрокліматом теплиці здійснювати за комплексним підходом, який охоплює комунікаційні та інформаційно-управляючі системи та

технології, сучасну елементну базу, програмне забезпечення з використанням ОС Android, засоби підтримки прийняття рішень і ґрунтується на таких принципах: системності, змінного складу обладнання, відкритості, модульності та використання комплексу базових проектних рішень. Мікроконтролер є центральним управляючим пристроєм всієї системи управління. Розроблено граф досяжності для комунікації між клієнтом та сервером для системи управління мікрокліматом теплиці, який має два кінцеві стани.

**Ключові слова:** мікроклімат, теплиця, управління, автоматична система управління, мікропроцесор, апаратне забезпечення, елементна база.

## Модель системы управления технологическим процессом в тепличном комплексе

А.В. Шарапа, А.Г. Бердников

**Шарапа Александр  
Вадимович**

*студент, студент 6-го курса кафедры теоретической и прикладной системотехники, Харьковский национальный университет имени В. Н. Каразина, площадь Свободы, 4, Харьков, Украина, 61022;*

**Бердников Анатолий  
Георгиевич**

*к.т.н., доцент; доцент кафедры теоретической и прикладной системотехники, Харьковский национальный университет имени В. Н. Каразина, площадь Свободы, 4, Харьков, Украина, 61022;*

Разработка модели автоматизированной системы управления микроклиматическими показателями и энергопотреблением тепличного хозяйства. Актуальной проблемой является проектирование интеллектуальных компонентов и синтез на их основе адаптивных компьютерных систем управления микроклиматом и энергоэффективностью тепличного хозяйства. Целью работы является анализ методов управления режимами работы тепличного хозяйства, сравнительная характеристика средств автоматизации управления технологическим процессом, разработка схемы информационного взаимодействия теплицы с окружающими и структурной схемы управления микроклиматом теплицы, подбор современных технических средств реализации систем автоматического регулирования микроклимата и исследования возможностей разработанной модели управления технологическим процессом. Методы исследований основываются на основных положениях теории тепло- и массообмена с использованием современных элементов компьютеризированного управления. Разработана структура системы управления микроклиматом в теплице, которая имеет переменный состав оборудования, обеспечивает быструю адаптацию к требованиям управления для конкретной теплицы. Предложено разрабатывать компоненты и синтезировать структуру системы управления микроклиматом теплицы осуществлять с комплексным подходом, который охватывает коммуникационные и информационно-управляющие системы и технологии, современную элементную базу, программное обеспечение с использованием ОС Android, средства поддержки принятия решений и основывается на следующих принципах: системности, переменного состав оборудования, открытости, модульности и использования комплекса базовых проектных решений. Микроконтроллер является центральным управляющим устройством всей системы управления. Разработан граф достижимости для коммуникации между клиентом и сервером для системы управления микроклиматом теплицы, имеющий два конечные состояния.

**Ключевые слова:** микроклимат, теплица, управления, автоматическая система управления, микропроцессор, аппаратное обеспечение, элементная база.

### Introduction

The agricultural sector of Ukraine is of great importance for the economy and is the basis of its citizens' well-being.

The agro-industrial sector of Ukraine formed about 13% gross domestic product (GDP) in 2018. The share of agricultural products in total Ukrainian export during this period amounted to 39.8%, or a record \$ 18.8 billion. The agricultural sector, due to its development, as well as, the decline of other industries, has established its status as a budget donor and among with information technologies (IT) is one of the leading sectors of the Ukrainian economy.

Since the greenhouses are an important part of the agricultural sector, the model of its management systems, its microclimatic indicators, as well as, its design and operational factors are highly relevant.

A greenhouse is an enterprise for the indoor growing of vegetables or seedlings for the open ground. The main structure of each greenhouse is a special cultivation room with a coating of translucent material (glass, polyethylene membrane). Heating in greenhouses can be provided by solar, biological (by biofuel), technical (hot water, steam, electricity, thermal waste of industrial enterprises, etc.) or geothermal energy sources. The main greenhouse crops are the vegetables (cucumbers, tomatoes,

peppers, lettuce, horseradish, parsley, radishes, onions, celery, cabbage), melons (zucchini, melons, watermelons), fruits (lemon, peach), berries (strawberries), decorative flowers (carnation, chrysanthemum, rose, etc.) and mushrooms.

Greenhouses are very cost-effective: the cost of their construction pays off in 4-5 years.

### **Automatic control system for process plan of a greenhouse**

Automatic control system for process plan (ACS PP) of a greenhouse is the basis of high yields and profitability of greenhouse production. It is known that the growth and development of plants directly depend on environmental conditions: the amount of light, heat, air quality, water, nutrients and other factors. Automated control of technological parameters allows us to control the microclimate in the greenhouse: temperature and humidity, carbon dioxide concentration, soil moisture and pH level, lighting, as well as remotely control the modes of greenhouse operation. The system of the automated microclimate control can include automatic watering, system of carbon dioxide replenishment, heating, ventilation, etc. Today, the ACS PP greenhouses are widely used by private farms to maintain the necessary temperature and humidity conditions in greenhouses.

The ACS PP of the greenhouses includes:

- automatic regulators, thermostats to maintain the optimum temperature;
- monitoring sensors to monitor temperature, CO<sub>2</sub> levels, humidity;
- methods of automatic control for measurement of the controlled parameters;
- annunciators to inform about the reaching the set limit values;
- power supplies, switching and protective equipment, manual controls;
- engineering systems: ground heating; electric illumination; irrigation system;
- complete automation cabinet, ventilation equipment, auxiliary components, etc.

Automation of the process of growing greenhouse products provides full conditions for the development and plant growth. The usage of hardware and software complex of automated control of technological parameters of greenhouses means increasing profits and lowering production costs. The advantages of ACS PP of the greenhouses appreciated by agronomists, technologists, farmers and gardeners, as automated process control allows increasing output of greenhouse products and profits. With the help of an intelligent automated system, it is possible to monitor the parameters in the greenhouse around the clock, and adjust them remotely if necessary.

### **Functional capabilities of ACS PP of the greenhouses**

"Smart greenhouse" is the key to successful crop production and economic efficiency of growing the greenhouse plants. ACS PP greenhouse operates as follows. The sensors record the values of technological parameters and transmit signals to the controller. Depending on the controller program the received data are processed and transferred to the personal computer or the smartphone in the form understandable by the operator (tables, mnemonics, graphics, etc.). Control signals can be issued from the controller to the activators to maintain the set values of technological parameters. By using the automated workplace the operator can set the necessary technological mode of operation of the greenhouse and control all ongoing processes.

Functionalities of ACS PP of the greenhouse:

- to control temperature automatically;
- to maintain the set parameters of a microclimate accurately;
- to display the technological information;
- to control and diagnostics of electrical equipment;
- to regulate process parameters.

ACS PP is a highly efficient tool for greenhouses, the purpose of which is to maintain optimal parameters for plant growth. Automation opens up new opportunities for greenhouse production and allows agribusiness to reach a new level. Thus, with the help of the automated climate control system of the greenhouse it is possible to maintain the required temperature and humidity regime, diagnose the technical condition of the equipment, use water, heat, energy resources economically. The introduction of ACS PP of the greenhouses helps to improve the quality of greenhouse products and increase the

efficiency of crop production by maintaining acceptable climatic conditions, minimizing losses and reducing plant illness.

### **Formulation of the problem**

The yield of greenhouses is largely determined by the microclimate, which is provided by modern computer control systems.

To evaluate the energy efficiency of greenhouses the indicator "energy consumption" is used, which characterizes the level of consumption of energy resources per unit of a product. The main problems of low energy efficiency of greenhouses are: excessive consumption of energy resources due to low energy efficiency of technologies and equipment and imperfection of schemes due to increased share of natural gas consumption and insufficient use of energy from alternative fuels and renewable sources; low level of management of energy efficiency and energy consumption.

Energy efficiency of greenhouses can be increased by such measures as the reduction of technological and non-production losses of energy resources by modernization of energy supply schemes and equipment, introduction of modern energy-efficient technologies, the optimization of the structure of consumption of fuel and energy resources, in particular, the replacement of traditional types of energy resources with other types such as renewable energy sources and alternative fuels, improvement of the management system of energy efficiency and energy consumption.

The main problem of developing intelligent components of microclimate control systems in greenhouses is the formation of requirements, the choice of data processing methods and the means of their implementation (software, hardware or software and hardware). Therefore, the actual problem is to design intelligent components and synthesize the adaptive computer systems for microclimate control and energy efficiency of greenhouses which adopt those components.

### **Software**

Modern human society is surrounded by a large number of auxiliary mechanisms that facilitate the life of the individual. Every year, technology is getting "smarter", taking over many routine tasks that people previously had to deal with.

The most basic things such as greenhouse management, traffic light regulation, supervising indoor ventilation or heating systems, controlling elevators, and millions of other daily routines previously required the direct involvement of a human operator. Now the technological devices are controlled by computers and microcontrollers such as the Arduino Mega 2560 board, the heart of which is the ATmega2560 microcontroller.

The microcontroller is represented by a RISC processor developed by AVR and operates at a frequency of 16 MHz, which is the greatest in the entire ATMel product line. Its chip contains all devices related to the general concept of a computer system: reprogramming static, as well as flash memory, interface bridges, multiplier.

The processor is characterized as a calculator of a single response time to execute any command, regardless of its complexity. The address bus and internal registers is 8 bits size. The maximum size of the external SRAM memory which can be connected is 64 KB. The frequency setting generator is part of the controller chip itself.

ATmega2560, in addition to standard, discrete parallel and serial inputs/outputs, is equipped with analog signal converters.

Also in addition to the ATmega2560 microcontroller, the Arduino Mega 2560 board includes the following components:

- USB connector;
- reset button;
- ICSP1 connector;
- LED indication;
- ICSP1 connector;
- voltage regulator - 3.3V;
- ATmega16U2 microcontroller;

- external power connector;
- voltage regulator - 5V.

### The main points of the model

The design of modern greenhouse climate control systems is based on system integration, which consists of a system approach covering all levels of integration of greenhouse management processes, and taking into account the requirements and efficiency of their application. It is expedient to develop greenhouse microclimate management systems on the basis of an integrated approach, which includes communication and information management technologies and systems, modern element base, software, decision support tools and is based on the following principles: consistency, equipment variability, openness, modularity, and usage of a set of basic design solutions.

The development of the modern control systems lies in the direction of reducing the size, weight and energy consumption, increasing reliability, functionality and intellectualization. Therefore, the new management systems based on new intelligent information technologies and a modern elemental base appear.

The microclimate in the greenhouse is controlled by a control system, the main components of which are: a microcontroller system, a mobile or PC operating system, sensors and activators (Fig. 1).

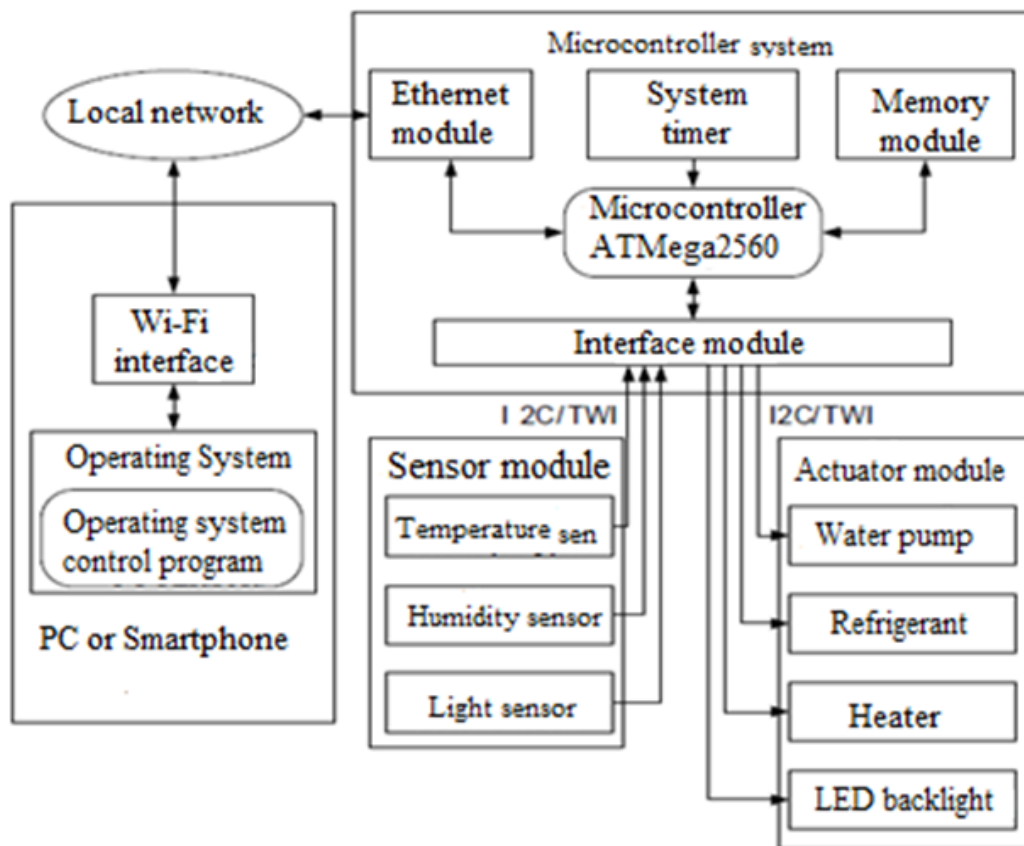


Figure 1. Block diagram of microclimate control in greenhouse

The microcontroller system consists of an Arduino board based on an ATmega2560 microcontroller, an Ethernet module, a memory module and an interface module.

The microcontroller connects all system modules into a single system. It is responsible for the autonomous management of the greenhouse microclimate. For this purpose, the microcontroller receives the sensor readings through the interface module and controls the activators according to the internal program.

The internal program analyzes the limits of the required parameters of the microclimate and generates control signals. If there is a connection to a mobile or computer control device, the

microcontroller checks for changes in the greenhouse control parameters and overwrites the configuration file.

By developing intelligent components and using them in the control system, it is possible to effectively manage the microclimate and energy efficiency of the greenhouse. The study of the operations proves increased efficiency of this model. After the analysis of the element base and other indicators, it has been found that the proposed Arduino board based on the ATmega2560 microcontroller is efficient and on a par with the competitors.

The accessibility graph for the client-server communication in the greenhouse microclimate control system has been constructed (Fig. 2)

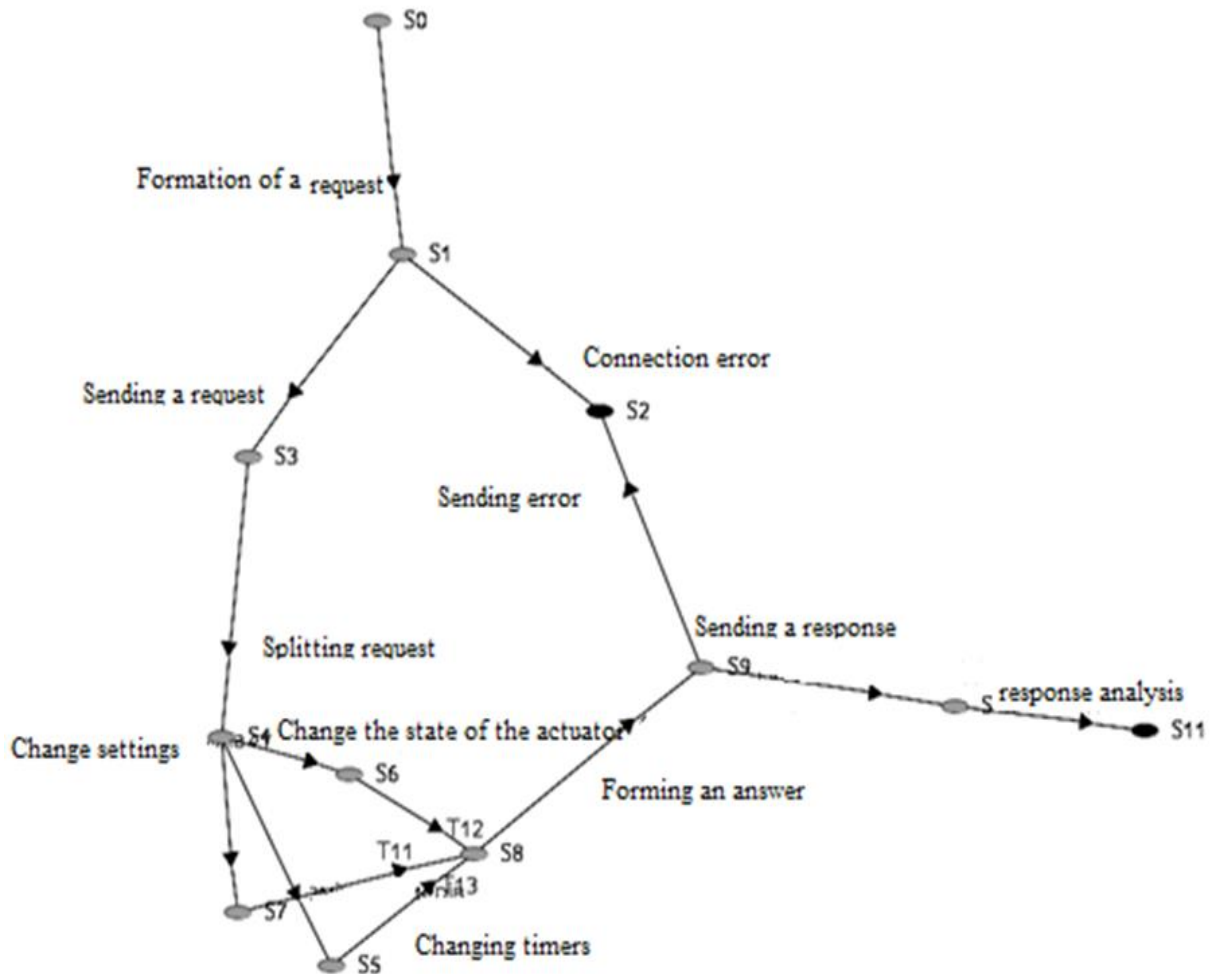


Figure 2. Accessibility graph for the client-server communication in the greenhouse microclimate control system

The accessibility graph for the client-server communication in the microclimate control system of a mini-greenhouse has two final states. The transition to the first final state means the successful communication between the client and server. The transition to the second final state means that a communication error has occurred and the requests have not been fulfilled. The accessibility graph is finite without blind spots and all states are reachable, therefore, the model works correctly.

### Conclusion

Important trends in modern agricultural production are the constant growth of its scale, increasing the quantity and quality of agricultural products, on the one hand, and the progressive shortage of labor, the unpopularity of monotonous and heavy manual labor in agriculture and animal husbandry on the other. The complex mechanization and automation of production are of the utmost importance and sometimes the single way of resolving those contradictions.



Thanks to mechanization and automation of production, labor productivity could be increased greatly. The implementation of integrated automation is of great economic importance, because it guarantees a positive economic effect.

Automation of agricultural production increases the reliability and prolongs the equipment life cycle, facilitates and improves working conditions, increases occupational safety and makes agricultural work more prestigious, reduces labor turnover and saves labor costs, increases the number and quality of products, accelerates the process of erasing boundaries between mental and physical, as well as, industrial and agricultural labor.

#### REFERENCES

1. V. V. Denisenko Computer control of technological process, experiment, equipment. - M.: Hot line Telecom, 2009. p.408. [in Russian]
2. Automation of technological processes and automatic control systems: Textbook / O.V. Baralo, P.G.Samoilenko, S.E. Granat, V.O. Kovalev - K.: Agricultural education, 2010. p. 3 - 4.
3. Agriculture of Ukraine, URL: <https://uk.wikipedia.org/>  
[https://uk.wikipedia.org/wiki/Сільське\\_господарство\\_України](https://uk.wikipedia.org/wiki/Сільське_господарство_України) [in Ukrainian]
4. Greenhouse: Geograf.com.ua URL: <http://www.geograf.com.ua/glossary/suspilno-geografichni-termyny/teplichne-gospodarstvo> [in Ukrainian]
5. Arduino Mega 2560: kosmodrom.com URL: <http://www.kosmodrom.com.ua/el.php?name=MEGA2560-R3-NO> [in Russian]
6. About the ATmega2560 controller, URL: <https://future2day.ru/pro-kontroller-atmega2560/> [in Russian]

#### ЛІТЕРАТУРА

1. Денисенко В. В. Компьютерное управление технологическим процессом, экспериментом, оборудованием. М.: Горячая линия Телеком, 2009. 408 с.
2. Автоматизація технологічних процесів і системи автоматичного керування: Навчальний посібник /Барало О.В., Самойленко П.Г., Гранат С.Є., Ковальов В.О. – К.: Аграрна освіта, 2010. – 3-4 с.
3. Сільське господарство України: <https://uk.wikipedia.org/> [Електронний ресурс] // Режим доступу: [https://uk.wikipedia.org/wiki/Сільське\\_господарство\\_України](https://uk.wikipedia.org/wiki/Сільське_господарство_України)
4. Тепличне господарство: Geograf.com.ua [Електронний ресурс] // Режим доступу: <http://www.geograf.com.ua/glossary/suspilno-geografichni-termyny/teplichne-gospodarstvo>
5. Arduino Mega 2560: kosmodrom.com [Електронний ресурс] // Режим доступу: <http://www.kosmodrom.com.ua/el.php?name=MEGA2560-R3-NO>
6. Про контроллер ATmega2560: Future2day.ru [Електронний ресурс] // Режим доступу: <https://future2day.ru/pro-kontroller-atmega2560/>



**ВІСНИК ХАРКІВСЬКОГО НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ  
імені В.Н. Каразіна**

Серія **«Математичне моделювання. Інформаційні технології. Автоматизовані системи управління»**

*Випуски даної серії розповсюджуються у академічних та наукових колах України та за її межами з метою оперативного висвітлення досліджень у таких актуальних галузях: математичне та комп'ютерне моделювання, обчислювальний експеримент, теорія і прикладні методи обробки інформації, захист інформації, програмно-апаратні системи інформаційного або управляючого призначення, застосування математичного моделювання та системного аналізу у високих, наукоємних технологіях, враховуючи технології створення програмної продукції. Приймаються роботи, що відносяться до напрямів фізико-математичних і технічних наук (бажаний об'єм 6-18 сторінок). Усі рукописи рецензуються.*

*Примітка. Протягом 2021-22 рр. редакційна колегія при інших рівних умовах надаватиме перевагу роботам, що представлені англійською мовою, якщо стаття отримала схвалення при рецензуванні.*

Офіційний сайт <http://periodicals.karazin.ua/mia>  
<http://mia.univer.kharkov.ua>

**Email: journal-mia@karazin.ua**

**ВЕСТНИК ХАРЬКОВСКОГО НАЦИОНАЛЬНОГО УНИВЕРСИТЕТА  
имени В.Н. Каразина**

Серія **«Математическое моделирование. Информационные технологии. Автоматизированные системы управления»**

*Випуски данной серии распространяются в академических и научных кругах Украины и за ее пределами с целью оперативного освещения исследований в следующих актуальных областях: математическое и компьютерное моделирование, вычислительный эксперимент, теория и прикладные методы обработки информации, защита информации, программно-аппаратные системы информационного или управляющего назначения, приложения математического моделирования и системного анализа в высоких, наукоёмких технологиях, включая технологии создания программной продукции. Принимаются работы, относящиеся к направлениям физико-математических и технических наук (желательный объем 6-18 страниц). Все рукописи рецензируются.*

*Примечание. На протяжении 2021-22 гг. редакционная коллегия при прочих равных условиях будет отдавать предпочтение работам, представленным на английском языке, если статья получила одобрение при рецензировании.*

Официальный сайт <http://periodicals.karazin.ua/mia>  
<http://mia.univer.kharkov.ua>

**Email: journal-mia@karazin.ua**

*Наукове видання*

**Вісник Харківського національного університету  
імені В. Н. Каразіна**

Серія «Математичне моделювання. Інформаційні технології.  
Автоматизовані системи управління»

Випуск 47

*Збірник наукових праць*

Українською, російською та англійською мовами

Комп'ютерне верстання О.О. Афанасьєва

Підписано до друку 28.09.2020 р.  
Формат 60x84/8. Папір офсетний. Друк цифровий.  
Ум. друк. арк. – 5,8.  
Обл.– вид. арк. – 7,3.  
Наклад 50 пр. Зам. № 66/2020  
Безкоштовно

Видавець і виготовлювач  
Харківський національний університет імені В. Н. Каразіна  
61022, м. Харків, майдан Свободи, 4  
Свідоцтво суб'єкта видавничої справи ДК №3367 від 13.01.09

Видавництво Харківський національний університет імені В. Н. Каразіна  
тел.: 705-24-32