The Role of Hungarian Traffic Rules Education and Examination System – a Quality Function Deployment Approach

Zsolt Csaba Horváth

Budapest University of Technology and Economics, Department for Railway Vehicles, Aircrafts and Ships, Műegyetem rkp. 3, H-1111 Budapest, Hungary, Email: horvath.zsolt@bmeits.hu

László Buics

Széchenyi István University, Department of Marketing and Management, Egyetem tér 1, H-9026 Győr, Hungary, Email: buics.laszlo@sze.hu

Péter Földesi

Széchenyi István University, Department of Logistics and Forwarding, Egyetem tér 1, H-9026 Győr, Hungary, Email: foldesi@sze.hu

Boglárka Eisinger Balassa

Széchenyi István University, Department of Marketing and Management, Egyetem tér 1, H-9026 Győr, Hungary, Email: eisingerne@sze.hu

Abstract: This paper examines the traffic rules education and examination system in Hungary, by using the Quality Function Deployment (QFD) method, as a new approach towards this complex topic. The education and examination of traffic rules are necessary for the stakeholders, but they have slightly different goals and objectives. This system has two separate stakeholders, the citizens and the authority, with their own set of goals, objectives, desires and ideas, about this system. The QFD reveals the connections between these layers. The paper analyses statistical data regarding road safety and presents the QFD model of both stakeholders and their inter-connections. The results of this work can be used to redesign education and examination methods, during the application of digitalized e-government solutions and as a general approach to match individual and public interests.

Keywords: Quality Function Deployment; Traffic rules; Traffic safety; E-government; E-learning

1 Introduction

In a perfect world, in a perfectly functioning state, all conditions and services are perfect. This means that citizens know and follow all the rules, the rules are perfectly designed, and citizens are physically and mentally able, for example, to learn and keep traffic rules perfectly. External conditions, such as the weather, are perfect and do not hinder this situation in any way. This situation only exists in an imaginary world, but the reality is far from that. In our study, we start from the suggestion of what is needed for road transport to work perfectly? This system has two separate stakeholders the citizens and the authority (Traffic Authority of Hungary), which is a governmental organization, representing the general interest of the public regarding this system. What can the citizen and the authority do to create this ideal situation? An approximately accurate answer to this question, can be obtained by examining the current system and detecting anomalies.

The relationship between the state and the citizen is twofold. The individual expects the state to protect him, so the citizen waives certain of his privilege and thus confers on the state, like enforcing regulations centralized. The purpose of the state is to protect the community; the purpose of the citizen is to protect itself. Therefore, the state must constantly examine whether the services it provides are appropriate for the citizen who confers power on him. If the state discovers problems or anomalies in the investigation of services, it must change them within the specified framework. Traffic rules can be different from country to country in terms of arrangements, over- and under-regulation. However, the main goal of the rules and regulations is not to overburden the citizens with legal knowledge, but to ensure the safety of all participants on the roads [6] [13] [16].

E-Government services are adding more and more service elements worldwide. Citizens have the privilege to conduct their affairs electronically, using the e-government system. The aim of the present study is to examine a part of e-government in Hungary, the online traffic rules education and examination system required to obtain a general driving license. During online traffic rules education, students can decide when and how to learn the curriculum and then take an examination at the end of the process. The online form of education is used in many countries around the world to learn traffic rules. In Hungary, the examination takes place electronically, but the candidate must appear at the examination venue personally [5] [7] [15].

In electronic government systems government operations are supported by web-based services. It involves the use of information technology, specifically the Internet, to facilitate the communication between the government and its citizens.

The citizen's satisfaction depends greatly on the level of the service and the quality of the product and the private companies do everything they can to acquire information about citizens needs in order to convert these expressed needs into new kind of products, shorter lead times, increased service levels [17]. The quality and efficiency of traffic rules education and examination is especially important in the age of developing autonomous vehicles in which case knowing and obeying traffic rules is crucial [18] [19].

This paper examines the traffic rule education and examination system in Hungary by using the Quality Function Deployment method as a new approach towards this complex topic, in order to be able to examine the different goals and objectives of the two separate stakeholders the citizens and the authority. Because of this the Quality Function Deployment method gives as a unique approach to examine these goals and objectives both separately and together to uncover the similarities and differences and their connection to each other [14] [20] [21] [22].

The paper analyses statistical data regarding road safety and accidents, classifies the goals and objectives of the citizens and the authority regarding traffic rules education and examination and provides a mathematical analysis with the help of the Quality Function Deployment approach presenting how the different goals and objectives are connected and what rules should be applied to enhance efficiency and effectiveness.

2 Statistical Analysis of Road Safety in Hungary

In this section the paper will show the statistical analysis of traffic safety. Accidents happen on the roads every day, a small fraction of these accidents are caused by mechanical or technical failures, others caused by drunkenness or other health issues. If we subtract these the main cause of the remaining accidents is mostly the insufficient knowledge of traffic rules or the non-compliance with the traffic rules.

If traffic rules were applied consciously and consistently the number of accidents could be decreased. With increased training and awareness, we assume that the number of accidents can be reduced. For this reason, we believe that great emphasis should be placed on education and application of the traffic rules, as we are of the opinion that in Central European countries only regular training and sanctions can enforce a positive improvement.

The number of road traffic accidents caused by drunk-driving (Figure 1) showed a slight decrease between 2010 and 2019. According to the data the proportion of road traffic accidents caused by drunk-driving between 2010 and 2019 can also be characterized by a slight decrease compared to other types of accidents, and the number of road accidents stagnated from 2016 to 2019. The number of fatal accidents also shows a declining trend (565 cases in 2016 and 530 in 2019).

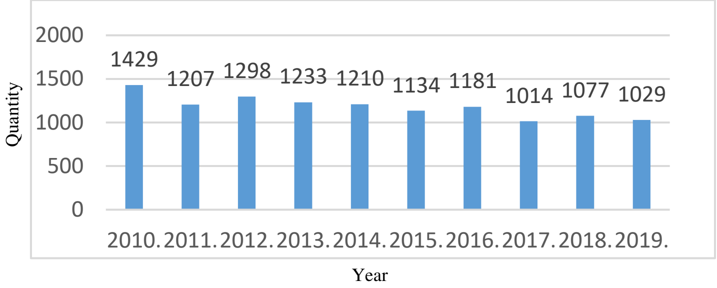


Figure 1

Number of road traffic accidents caused by drunkenness (2010-2019)

Source: Authors' own creation based on Hungarian Central Statistical Office data

Personal injury road accidents show a slight increase in the categories of easily and severely injured between 2016 and 2019, while the number of fatalities decreased slightly (2016: 607 cases, 2019: 602 cases). The number of personal injury road accidents (Figure 2) is mostly caused by passenger cars, their number increased between 2016 (10606 cases) and 2019 (10865 cases), and in 2018 there was a high number of cases (10920). The second most common cause is accidents caused by bicycles and the third most common is caused by accidents involving a freight vehicle. Numbers of 2020 are extrapolated based on the available data.

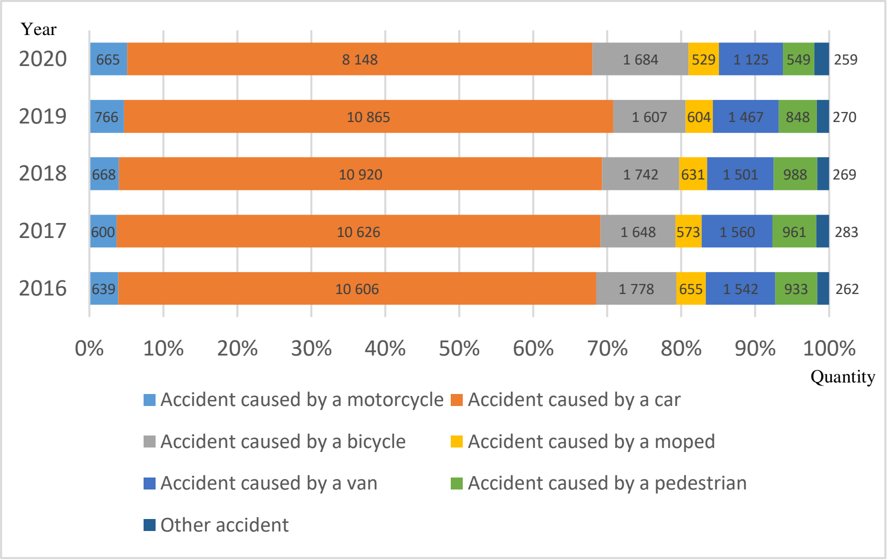


Figure 2

Number of road traffic accidents with personal injuries by causes (case)

Source: Authors' own creation based on Hungarian Central Statistical Office data

Figure 3 shows the causes of the accidents. Numbers of 2020, are extrapolated based on the available data. The most common cause is road track failure and other causes, the number of cases of which has increased slightly since 2016. The second most common cause is an accident due to pedestrian fault, but their number decreased slightly between 2016 and 2019.

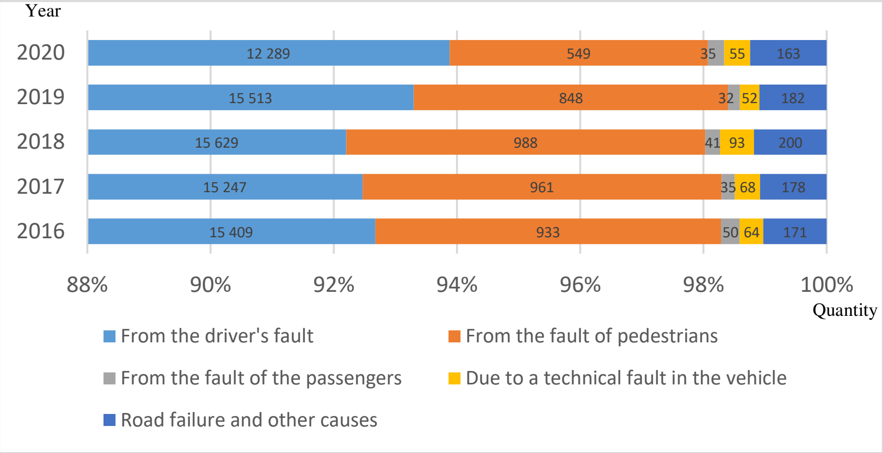


Figure 3

The causes of the accident (cases)

Source: Authors' own creation based on Hungarian Central Statistical Office data

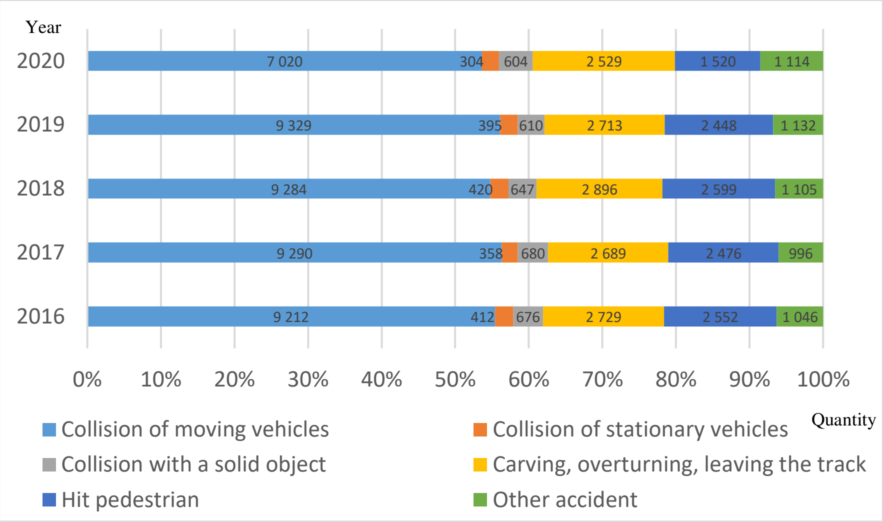


Figure 4

Number of road accidents with personal injuries by nature of the accident (case)

Source: Authors' own creation based on Hungarian Central Statistical Office data

The number of personal injury road accidents is shown by type in Figure 4. The most common are collisions with advanced vehicles, the number of these cases increased between 2016 and 2019 (from 9212 cases to 9329 in the evening).

The second most common accident is carving, overturning, leaving the track and the third is hitting a pedestrian.

Personal injuries are most often caused by drivers' faults. This number of cases did not change significantly between 2001 and 2019, with some decrease. It was exceptionally low in 2012 and exceptionally high in 2006. The number of road accidents caused by pedestrians decreased during the period under review. The results for 2020 are extrapolated, based on the available data.

3 Research Concept

Our research model is described in Figure 5. The areas examined in our study are the e-learning system that can be used to master the traffic rules and the examination system used for the traffic rules examination. As we examine two stakeholders (citizen, authority), we included both stakeholders and both areas in our model. In this study we use the QFD model and we create Quality Function Deployment matrixes [8] [11] [12].

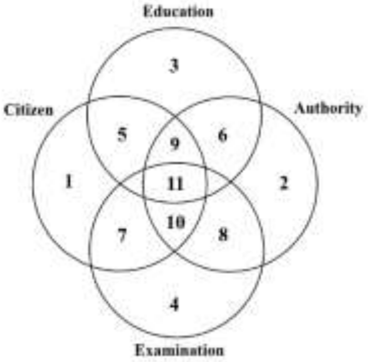


Figure 5
Research model
Source: Authors' own creation

In Figure 1 the numerical labels represent the different sections where the interests of the two stakeholders meet and overlap the education and the examination systems and each other on different levels. Respectively for example the segment labeled by 1 only represent the citizens, the segment labeled by 5 include only citizens and the education system while the segment labeled by 9 include citizens, education system and authority as well.

The research model is based on the contradiction that the citizen's and the authority's goals are not aligned with each other. It is in the citizen's interest to obtain a valid driving license, as that document allows to avoid being penalized while driving. For the authority, the existence of the license gives the strong assumption that the citizen knows the traffic rules. (In the present study, we deal with the theoretical examination of the traffic rules, we do not cover practical learning and the examination). Thus, while from the point of view of the authority ensuring the necessary knowledge is the final goal, for the citizens the knowledge is just a necessity to be able to drive without penalty.

In this research model, we created two Quality Function Deployment matrixes one for the e-learning system and the other for the examination (in both cases, we examined the perspectives of the citizen and the authority on services separately).

Then, to explore the anomalies and similarities, we use the method - already known and often used in the QFD methodology, - when we identify the opinions and subjective perceptions of the two stakeholders by “stacking” them, meaning that the different preferences of the two stakeholders were combined into one system

In Figure 1, showing the research model, the first number (1) refers to the citizen whose external and internal characteristics influence the learning process and the success of the examination, whose basic purpose is to obtain a license. Number two (2) means the authority whose purpose is to ensure that the driving citizen is licensed, knows the rules of the road and does not cause an accident. Number three (3) is the e-learning system. Number four (4) is the traffic rule examination. The QFD methodology was used in each case to examine further correlations. Number five (5) presents the evaluation of e-learning education from the citizen's perspective. Number six (6) is from the authority's point of view of traffic evaluation education. Number seven (7) points out the anomalies and similarities between the citizen and the authority in the evaluation of road traffic education.

Number eight (8) presents the evaluation of the traffic examination from the citizen's point of view. Number nine (9) shows from the authority's point of view of the traffic test assessment. Number ten (10) points out the anomalies and similarities between the citizen and the authority in the evaluation of the traffic rule examination. The advantage of the model outlined above is that it is able to examine all stakeholders and the activities that arise during the successful acquisition of the theoretical test of the license, to identify areas for the improvement of services.

If we look at the issue of road traffic e-learning education and examinations, the question is: how does the state change the system to make citizens be satisfied? To explain the topic, we take into account accident statistics that accurately show the various causes of accidents, the severity of personal injuries, or accidents by their nature. These statistics are discussed later in this study. The other aspect is the examination of traffic education and examination.

4 Design Procedure for Examining Stakeholder Connections

In our research, we use the Quality Function Deployment approach [1] [2] to structurally compare the goals and objectives of the of the citizens and the authority, both in case of the traffic rules education and the examination process, provided by e-government solutions. According to the literature the QFD methodology is a flexible design tool for products and services. In this context we apply the QFD in order to unravel the nature of connections between the goals and tools of the two stakeholders. However, in contrast to other products, where the interests of the stakeholders are not so different, in case of this e-government service much more complex. The learner-driver (citizen) wants to get the driving license as easy as possible. The traffic authority wants to achieve complete compliance with the traffic rules to minimize accidents. Because of this, we create two different Quality Function Deployment matrixes, combining the goals and objectives of the citizens and the authority.

During the application of the QFD model [3] [10] we formulated the following research questions:

1) What are the expectations of the citizens and the authority in connection with the driving license examinations?

2) What discrepancies can be identified between the two stakeholder groups?

Our research goal is to analyze the expectation of two stakeholders' group, and identify discrepancies, so we have to build one Quality Function Deployment matrix [4] [9] [24] for citizens one for authority.

4.1 Summary of Education and Examination Goals and Objectives of Citizens and the Authority

We present the expectations of the citizens for the two target groups and for the education and examinations separately. Based on our research model, we examine the sections marked with numbers 5-6-7-8 (Figure 1). We apply the QFD method for these sections, by combining the goals and objectives of citizens and the authority regarding the education and examination of traffic rules. The examination of the similarities and differences between the QFDs is examined by the target groups. Thus, we identify based on focus group interviews (two focus groups with 16 participants) the goals and objectives of citizens and the authority for both traffic rule learning and examinations. These goals and objectives serve as tools for the methodological study, to show how the QFD can be used to match the individual and public interest. In the future, it is planned to make a more detailed quantitative research to identify other objectives and goals.

Table 1 summarizes the goals and objectives of the citizens and the authority towards the education of traffic rules while Table 2 summarizes the goals and objectives towards the examination. The goals and objectives were defined based on expert opinions to present the application and main process steps of QFD and to highlight its usefulness.

The initial list of goals and objectives can be expanded and reformed during future researches. According to the principles of E-government in order to achieve a higher level of efficiency the appropriate tools should be applied, thus the authors' goal is to use the current list of goals and objectives to demonstrate the usefulness of the Quality Function Deployment method in this current context.

Table 1
Quality Function Deployment: Education
Source: Authors' own creation

Education			
GOALS	**Citizens**	**G1**	It should be easy to handle
		G2	It should be quick
		G3	Offline usability
		G4	Have smaller modules
		G5	Use simple and clear examples
		G6	Good quality illustrations and animation
		G7	Be up to date
		G8	Be free
	Authority	**G9**	Skill level application of traffic rules
		G10	Continuous knowledge control of traffic rules
		G11	Minimizing education and maintenance costs
OBJECTIVES	**Citizens**	**O1**	It should be easy to use on any device
		O2	Be accessible anywhere
		O3	Whatever time is available for learning
		O4	There should be more market players
		O5	Possibility of pre-trial
		O6	All citizens should receive state support
		O7	All students should be given a digital tool for learning
	Authority	**O8**	Continuous updating and development of system and knowledge material
		O9	Continuous monitoring and feedback function during training
		O10	Education can be solved without the use of human resources

Table 1 contains the goals and objectives of the citizens and the authority based on the research. Goals are general aims of the service design which pave the way of development in order to improve the overall quality. Objectives are more specific issues which have to be addressed in order to improve overall quality. They can be connected to one or several goals either in a positive or negative way, and there can be also goals to which they are not connected at all.

It is possible to create groups in Table 1 within the goals and objectives of the citizens and the authority, for example in case of education G1, G2 and G3 goals can be all connected to the User-friendly design, G4, G5, G6 goals can be connected to the easy learning. G9 and G10 goals can be connected to road safety. Regarding the objectives of education O1, O2 and O3 objectives can be connected to accessibility, O4 and O5 can be connected to the freedom of choice regarding the learning systems. Regarding the education citizens would prefer a user friendly and flexible environment with easy accessibility to the learning materials, and more opportunity to test their knowledge before the take the examination. Regarding the education, the authority's main goal is to ensure the safety on the roads by demanding high level of knowledge regarding the rules and regulations of traffic participation.

Table 2
Quality Function Deployment: Examination
Source: Authors' own creation

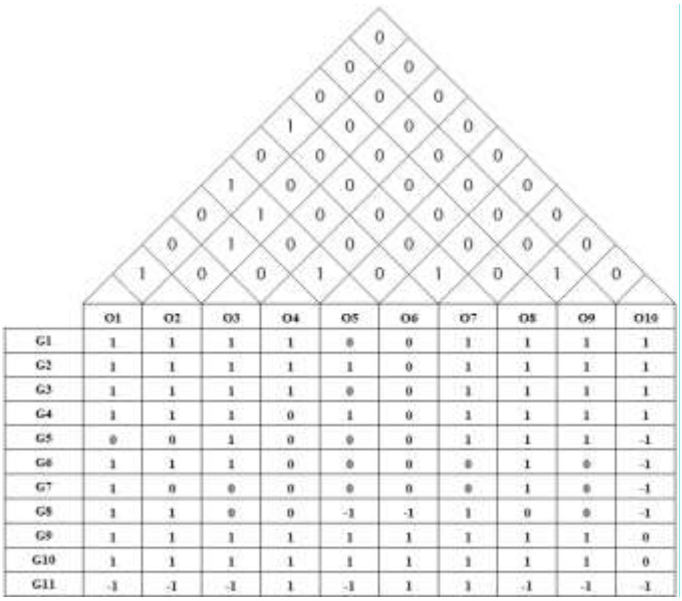
Examination			
GOALS	**Citizens**	**G1**	Pass the examination the first time
		G2	Use simple and clear examples
		G3	Make the examination system easy to use
		G4	Flexible examination times
		G5	The examination location should be easily accessible
	Authority	**G6**	Minimization of accidents due to violation of rules
		G7	100% of all licensed learn the rules of the road.
		G8	Automation of the entire examination system
		G9	Maintain the current course of the examination
OBJECTIVES	**Citizens**	**O1**	Examination dates that are more flexible to the client's needs
		O2	The examination can be taken at any time
		O3	Online examination opportunity
		O4	Multiple examination locations
	Authority	**O5**	Introduction of a system of proficiency tests
		O6	Collect as many fees as possible from participants
		O7	The examination points should remain in their current form
		O8	The need for human resources should remain in the process

It is possible to create groups in Table 2 within the goals as well. G2, G3 goals can be connected to the content of the examination, G4 and G5 can be connected to the organization of the examination. Regarding the examination citizens' main goal is to pass the examination in order to acquire the driving license, which is necessary to participate legally in the road traffic. In order to do so they would prefer a more flexible examination environment with an easily usable examination system and well-designed content. Regarding the examination, the main goal of the authority is to minimize the traffic accidents by demanding high level of knowledge regarding the rules and regulations of traffic participation.

4.2 Connections of Goals and Tools and Suggested Future Development Conditions

During the application of the Quality Function Deployment methodology we compared the goals of citizens and the authority regarding the education and the examination against the objectives of the two stakeholders. The results can be seen in Figures 6 and 7, where we used the mathematical formulas represented in Figure 8 to numerically express whether there is a positive, negative or neutral connection between the goals and objectives of the stakeholders and also between the objectives themselves. According to the Quality Function Deployment method objectives can be interpreted as tools serving the goals (desires) of the stakeholders. During a multiple level examination, these tools can be interpreted as goals of a lover level. Because of this high flexibility on some occasions there are no sharp borders between the content of goals and tool on the same level either.

$$r_{ij} = \begin{cases} 1, & \textit{positive connection} \\ 0, & \textit{neutral connection} \\ -1, & \textit{negative connection} \end{cases}$$

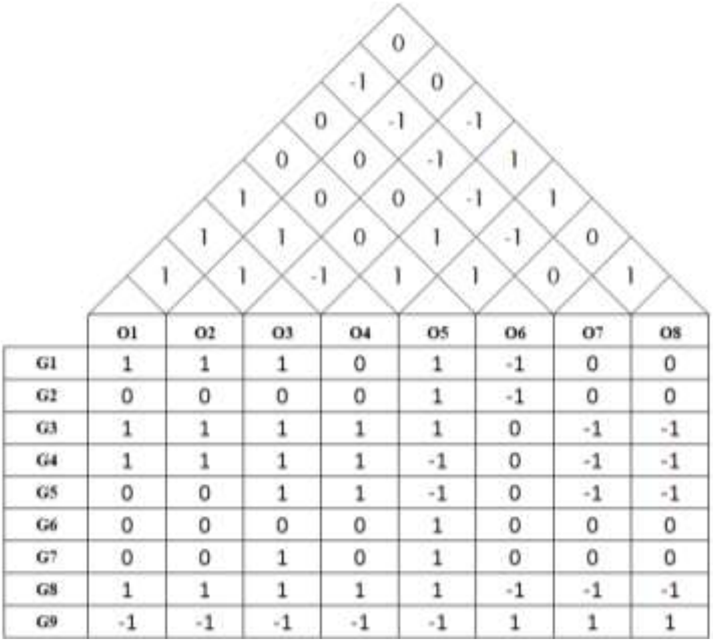

0
0 0
0 0 0
1 0 0 0
0 0 0 0 0
1 0 0 0 0 0
0 1 0 0 0 0 0
0 1 0 0 0 0 0 0
1 0 0 1 0 1 0 1 0

	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
G1	1	1	1	1	0	0	1	1	1	1
G2	1	1	1	1	1	0	1	1	1	1
G3	1	1	1	1	0	0	1	1	1	1
G4	1	1	1	0	1	0	1	1	1	1
G5	0	0	1	0	0	0	1	1	1	-1
G6	1	1	1	0	0	0	0	1	0	-1
G7	1	0	0	0	0	0	0	1	0	-1
G8	1	1	0	0	-1	-1	1	0	0	-1
G9	1	1	1	1	1	1	1	1	1	0
G10	1	1	1	1	1	1	1	1	1	0
G11	-1	-1	-1	1	-1	1	1	-1	-1	-1

Figure 6
Education QFD of goals and objectives
Source: Authors' own creation

As we can see of Figure 6 regarding education most of the objectives of the citizens and the authorities are neither in a positive nor in a negative connection with each other which indicates that the two stakeholders have a mostly different mindset regarding the subject.

As we can see on Figure 7 regarding the examination there is a more negative connection between the different stakeholders' objectives indicating that the citizens and the authority's expectations are in a stronger contrast with each other regarding this subject.



	O1	O2	O3	O4	O5	O6	O7	O8
G1	1	1	1	0	1	-1	0	0
G2	0	0	0	0	1	-1	0	0
G3	1	1	1	1	1	0	-1	-1
G4	1	1	1	1	-1	0	-1	-1
G5	0	0	1	1	-1	0	-1	-1
G6	0	0	0	0	1	0	0	0
G7	0	0	1	0	1	0	0	0
G8	1	1	1	1	1	-1	-1	-1
G9	-1	-1	-1	-1	-1	1	1	1

Figure 7
Examination QFD of goals and objectives
Source: Authors' own creation

Figure 8 summarizes the matrixes and vectors describing the QFDs regarding the Education and Examination goals and objectives of the citizens and the authority. In this context we are using matrix calculations to describe the connection between the variables. As we can see in Figure 8, we make a difference between the ranking of goals and objectives according to citizens and authority giving us further insights into the details. We also include the costs of the different objectives, as a variable, in our formulas, to help making difference between the resource requirements of the objectives.

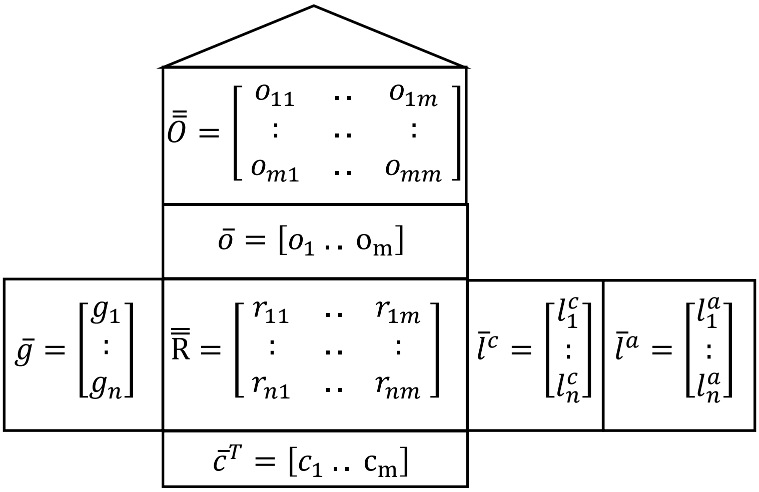

$\bar{\bar{O}} = \begin{bmatrix} o_{11} & .. & o_{1m} \\ \vdots & .. & \vdots \\ o_{m1} & .. & o_{mm} \end{bmatrix}$
$\bar{o} = [o_1 .. \mathrm{o_m}]$
$\bar{g} = \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix}$
$\bar{\bar{\mathrm{R}}} = \begin{bmatrix} r_{11} & .. & r_{1m} \\ \vdots & .. & \vdots \\ r_{n1} & .. & r_{nm} \end{bmatrix}$
$\bar{l}^c = \begin{bmatrix} l_1^c \\ \vdots \\ l_n^c \end{bmatrix}$
$\bar{l}^a = \begin{bmatrix} l_1^a \\ \vdots \\ l_n^a \end{bmatrix}$
$\bar{c}^T = [c_1 .. \mathrm{c_m}]$

Figure 8

Mathematical formulas for analysis and development

Source: Authors' own creation

$\bar{g} = Goals\ of\ citizens\ and\ authority$

$\bar{o} = Objectives\ of\ citizens\ and\ authority$

$\bar{\bar{O}} = o_{n1} \cdot o_{1n} = \bar{\bar{O}}_{nn}$ *Cross-connection between objectives*

$\bar{\bar{\mathrm{R}}} =$ *Connection between goals and objectives*

$$r_{ij} = \begin{cases} 1, & positive\ connection \\ 0, & neutral\ connection \\ -1, & negative\ connection \end{cases}$$

$\bar{c}^T = Costs\ of\ objectives$

$\bar{l}^c = Ranking\ of\ goals\ according\ to\ citizens$

$\bar{l}^a = Ranking\ of\ goals\ according\ to\ authority$

In the following, we present six conditions (A, B, C, D, E, F) which can be used to evaluate the connections between the goals and objectives of the stakeholders and the cross correlation of objectives, uncovering weaknesses and opportunities of development.

(A) condition

According to the first condition, we state that each objective should have more benefits than harm.

$$r_{ij} \in \{-1; 0; 1\} \qquad (1)$$

$$\sum_{i=1}^{n} r_{ij} > 0\ \ \forall\, j = 1 \ldots m \qquad (2)$$

$$r_{ij} = \begin{cases} 1, & positive\ connection \\ 0, & neutral\ connection \\ -1, & negative\ connection \end{cases} \qquad (3)$$

(B) condition

According to the second condition, we state that each tool should have more benefits than harm by including into the formula the weights of the different goals as well.

$$\sum_{i=1}^{n} r_{ij} \cdot l_i > 0 \ \forall j = 1 \dots m \qquad (4)$$

Weights represent the order of importance given by the citizens and the authority regarding the goals, but more research is needed to define these weights properly. In this paper we only present an example of usage, presenting that stakeholder priorities are different and depending on the priorities of the legislative act, it will be different if the priorities change.

(C) condition

According to the third condition we state that all targets should be supported to a greater than zero extent, which means that all tools should have a positive role and there should be at least one supporting tool for each purpose, so it must be more than zero when summed line by line.

$$\sum_{j=1}^{m} r_{ij} > 0 \qquad i = 1 \dots n \qquad (5)$$

(D) condition

According to the fourth condition we state that all objectives (tools) should have a supporting role in the design, thus after the categorization of their roles we can determine which of them should be changed or excluded from the design.

$$o_{ij} \in \{-1; 0; 1; -\infty\} \qquad (6)$$

$$o_{ij} = o_{ji} \qquad (7)$$

$$o_{ii} = 0 \qquad (8)$$

$$o_{ij} \neq -\infty \qquad (9)$$

$$\sum_{i=1}^{n} o_{ij} > 0 \ \forall j = 1 \dots m \qquad (10)$$

$$If\ o_{ij} \begin{cases} > 1, & \text{support} \\ = 0, & \text{neutral} \\ < -1, & \text{weaken} \\ = -\infty, & \text{antagonistic} \end{cases} \qquad (11)$$

According to our definition an antagonistic connection can also be presumed, between the objectives, indicated by the (9) formula. In this case, the objectives are incompatible, thus, one of them should be removed or redesigned.

This analysis also requires more data collection and research but this paper provides the initial concept, which can be further tailored by also adding sensibility and coherency tests.

Figures 9 and 10 presents an example regarding the first four conditions in case of education, showing that based on our initial results currently not every goal is supported to a greater than zero extent and there are tools with more negative than positive effects.

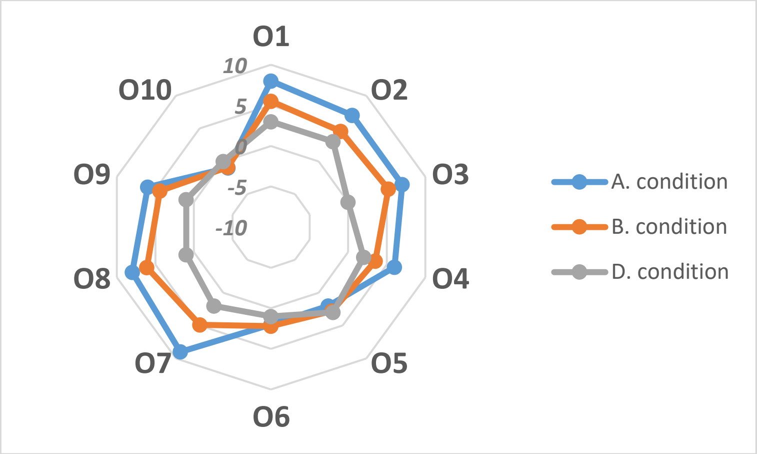


Figure 9

Education objectives according to A., B., and D. conditions

Source: Authors' own creation

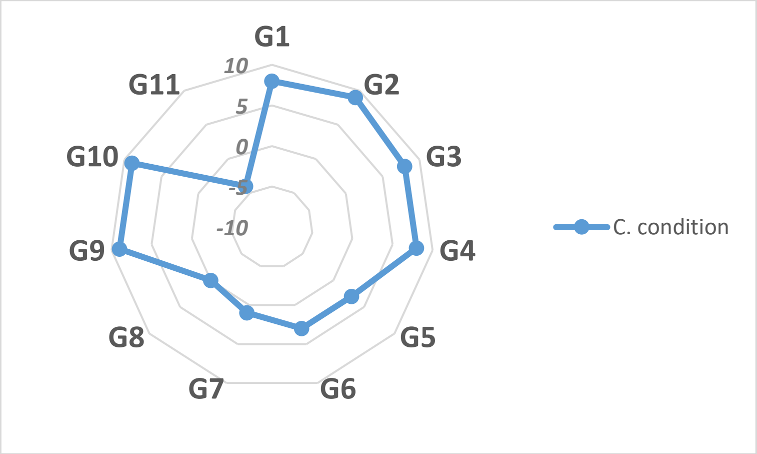


Figure 10

Education goals according to C. condition

Source: Authors' own creation

Figures 11 and 12 presents an example regarding the first four conditions in case of examination, showing that based on our initial results currently not every goal is supported to a greater than zero extent and there are tools with more negative than positive effects as well.

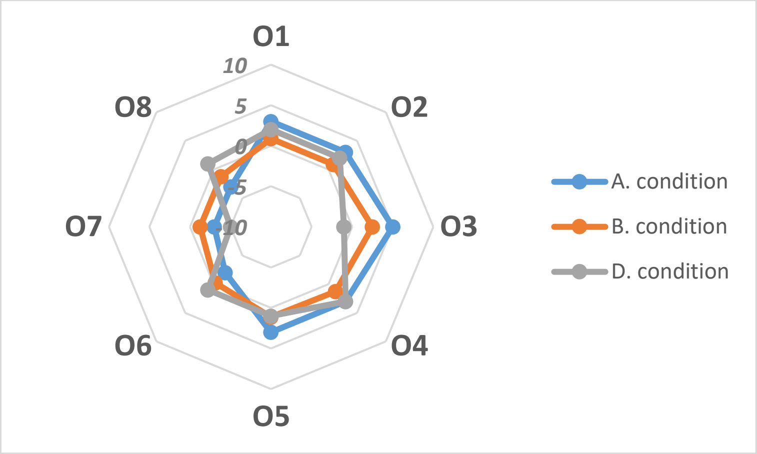


Figure 11
Examination objectives according to A., B., and D. conditions
Source: Authors' own creation

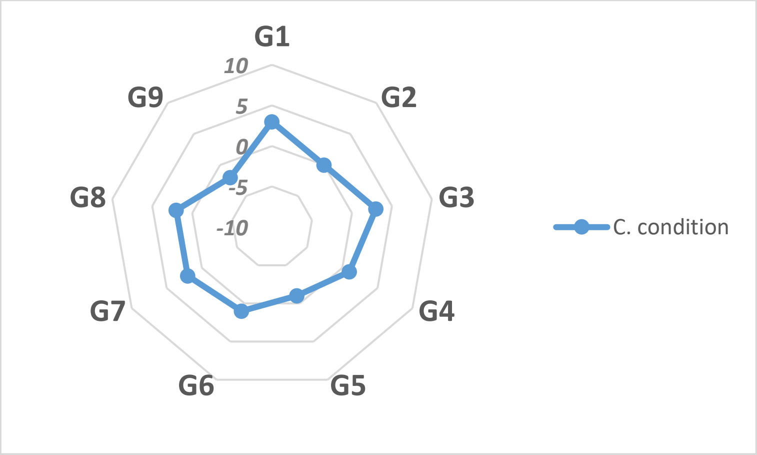


Figure 12
Examination goals according to C. condition
Source: Authors' own creation

The total benefit of each asset can be calculated by using the following formula after determining cost values to each objectives (tools) after a more detailed analysis. This calculation could also help to determine which tools should be

(12)

changed or excluded from the design, after properly defining the individual costs of the objectives.

$$\frac{\sum_{i=1}^{n} r_{ij} \cdot l_{ij}}{C_j} > 0 \ \forall j = 1 \ldots n$$

(E) condition

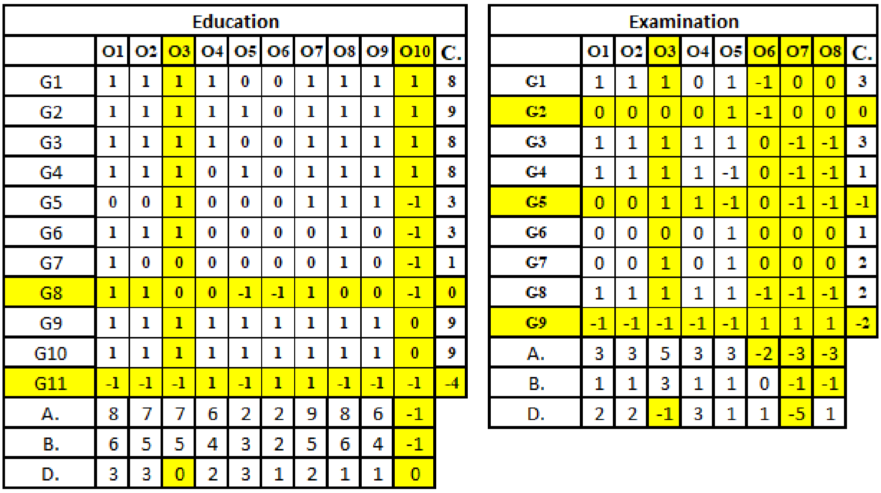
For further development we suggest the application of the fifth condition which ensures that all objectives should have a supporting effect in the design and innovation of the education and the examination system, tools having neutral or negative effects should be excluded entirely.

(F) condition

Last but not least by applying the sixth condition we suggest that general coherence and consistency test should be executed during the development to enhance efficiency and regardless of the sum, the number of pieces with negative signs cannot be more than (m-1) / 2

$$Number\ of\ negative\ elements \leq Z\left(\frac{m-e}{2}\right) \tag{13}$$

e: efficiency



Education

	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10	C.
G1	1	1	1	1	0	0	1	1	1	1	8
G2	1	1	1	1	1	0	1	1	1	1	9
G3	1	1	1	1	0	0	1	1	1	1	8
G4	1	1	1	0	1	0	1	1	1	1	8
G5	0	0	1	0	0	0	1	1	1	-1	3
G6	1	1	1	0	0	0	0	1	0	-1	3
G7	1	0	0	0	0	0	0	1	0	-1	1
G8	1	1	0	0	-1	-1	1	0	0	-1	0
G9	1	1	1	1	1	1	1	1	1	0	9
G10	1	1	1	1	1	1	1	1	1	0	9
G11	-1	-1	-1	1	-1	1	1	-1	-1	-1	-4
A.	8	7	7	6	2	2	9	8	6	-1	
B.	6	5	5	4	3	2	5	6	4	-1	
D.	3	3	0	2	3	1	2	1	1	0	

Examination

	O1	O2	O3	O4	O5	O6	O7	O8	C.
G1	1	1	1	0	1	-1	0	0	3
G2	0	0	0	0	1	-1	0	0	0
G3	1	1	1	1	1	0	-1	-1	3
G4	1	1	1	1	-1	0	-1	-1	1
G5	0	0	1	1	-1	0	-1	-1	-1
G6	0	0	0	0	1	0	0	0	1
G7	0	0	1	0	1	0	0	0	2
G8	1	1	1	1	1	-1	-1	-1	2
G9	-1	-1	-1	-1	-1	1	1	1	-2
A.	3	3	5	3	3	-2	-3	-3	
B.	1	1	3	1	1	0	-1	-1	
D.	2	2	-1	3	1	1	-5	1	

Figure 13

Education and Examination goals and objectives which do not meet a defined condition

Source: Authors' own creation

Figure 13 summarizes which goals and objectives do not meet with one or more defined conditions regarding the presented example. As we can see in case of the education both under supported goals are connected to the cost of education systems and materials, while in case of the objectives the flexibility and the human resource necessity fail to meet the conditions. Under examination, we can see several goals and objectives require reconsideration. The troublesome goals are in connection with the flexibility of the examination system and examination content, while in the case of the objectives, again the flexibility and the human resource necessity, fail to meet the conditions set.

Conclusions

This paper focused on the goals and objectives of both the citizens and the authority, regarding traffic rules education and examination. The work employed the model of Quality Function Deployment (QFD), to compare these goals and objectives with each other, highlighting both positive and negative connections between them, while also providing a thorough statistical background analysis, regarding the topic of road safety.

Our paper suggests a set of measures, as a design and control technique, for the evaluation of the current connections between the goals and objectives of the stakeholders and also for helping during the design and development of these systems.

The relationship between the authority and the citizen is twofold. The authority wants to ensure that every citizen knows and abides by the traffic rules, in order to minimize the number of accidents on the roads, while the citizens' major goal is to be able to legally use vehicles on the roads and learning the rules is just a means towards that goal.

The suggested conditions can help to parameterize the elements of the design. As we start to tighten the framework, with each condition, it can help to determine which tool should be reconsidered or even excluded from the design, in order to enhance the efficiency of the system, for the sake of both stakeholders.

This paper presents our initial results, with examples of the application of the conditions, which will be further tailored, after a more detailed survey and statistical research is carried out. In our paper, we determined that these methods are useful and can be applied to this complex issue, using the Quality Function Deployment approach, as we are able to identify key points, where improvements can be made. In our future research, we will put a greater focus on the connections and rankings, by collecting more detailed data, using analytical methods, in order to point out inefficiencies, which are worth further research and examination.

Acknowledgement

The research presented in this paper was funded by the "Thematic Excellence Program – National Challenges Subprogram – Establishment of the Center of Excellence for Autonomous Transport Systems at Széchenyi István University (TKP2020-NKA-14)" project.

References

[1] Adiandari, A., Winata, H., Fitriandari, M., & Hariguna, T. (2020) Improving the quality of Internet banking services: An implementation of the quality function deployment (QFD) concept. Management Science Letters, 10(5), pp. 1121-1128

[2] Akao, Y. (1990) Quality Function Deployment: Integrating Customer Requirements into Product Design; Productivity Press: Cambridge, MA, USA, pp. 25-50

[3] Alinizzi, M., Haider, H., Almoshaogeh, M., Alharbi, F., Alogla, S. M., & Al-Saadi, G. A. (2020) Sustainability Assessment of Construction Technologies for Large Pipelines on Urban Highways: Scenario Analysis Using Fuzzy QFD. Sustainability, 12(7), p. 2648

[4] Bagassi, S., De Crescenzio, F., Piastra, S. (2020) "Augmented reality technology selection based on integrated QFD-AHP model", International Journal on Interactive Design and Manufacturing (IJIDeM), Volume 14, pp. 285-294, https://doi.org/10.1007/s12008-019-00583-6

[5] Buics, L., Eisinger Balassa, B. (2020) "Applying new methods for analyzing public service processes", Pro Publico Bono, Vol. 8, No. 2, 8(2), pp. 2-29

[6] Buics, L., Eisingerné Balassa, B. (2020) Servitization of public service processes with a simulation modelling approach, Engineering Management in Production and Services, 12(3), pp. 116-131

[7] Buics, L., Süle, E. (2020) Statistical analysis of Hungarian public service processes for key performance indicator measurement. Hungarian Statistical Review, 3 (2) pp. 71-98

[8] Cohen, L. (1995) "Quality Function Deployment: How to Make QFD Work for You", Addison- Wesley, Reading, MA. pp. 28-98

[9] David, L., Goetsch, Davis, S. (2000) "Quality management: introduction to total quality management for production, processing, and services", Prentice Hall, pp. 20-65

[10] Dejanović, A., Nikolic, S., Stankovic, J. (2015) Integral Model of Strategic Management: Identification of Potential Synergies. Acta Polytechnica Hungarica. 12, pp. 115-133

[11] Hauser, J. R., Griffin, A., Klein, R. L., Katz, G. M., Gaskin, S. P. (2010) "Quality function deployment (QFD)", Wiley International Encyclopedia of Marketing. pp. 1-16

[12] Hauser, J.; Clausing, D. (1988) "The House of Quality. Harv. Bus. Rev.",Volume 66, pp. 63-73

[13] Hood, C., Peters, G. (2004) The Middle Aging of New Public Management: Into the Age of Paradox? Journal of Public Administration Research and Theory: J-PART, Vol. 14, No. 3 (Jul., 2004), pp. 267-282

[14] Jeong C. H. I. (2007) Fundamental of Development Administration, Scholar Press,, Selangor, pp. 87-135

[15] Kolnhofer-Derecskei; A., Zsuzsánna Reicher, R., Szeghegyi, Á., (2019) "Transport habits and preferences by generations - does it matter regarding the state of the art? ", Acta Polytechnica Hungarica, pp. 29-44

[16] Machta, J. (1999) Entropy, information, and computation. American Journal of Physics 67, p. 1074

[17] Olewnik, A., Lewis, K. (2008) Limitations of the House of Quality to provide quantitative design information. International Journal of Quality &amp Reliability Management, 25, pp. 125-146

[18] Péter, T., & Lakatos, I. (2019) Vehicle dynamic-based approach for the optimization of traffic parameters of the intelligent driver model (IDM) and for the support of autonomous vehicles' driving ability. Acta Polytechnica Hungarica, 16(3), pp. 121-140

[19] Prusty, S., Mohapatra, P., Mukherjee, C. (2017) House of Strategy: A model for designing strategies using stakeholders' opinion. Computers & Industrial Engineering. 108, pp. 39-56

[20] Suhardi, A. R. (2013, October) Quality Function Deployment to Improve Quality of Service. The 2nd 2013 IBSM International Conference on Business and Management, Trisakti University Indonesia, Prince of Songkla University thailand, Budi Luhur University Indonesia

[21] Sullivan, L. P. (1986) "Quality function deployment. Quality Progress", pp. 39-50

[22] Yamamoto, C., Kishi, K., Hara, F., & Satoh, K. (2005) Using quality function deployment to evaluate government services from the customer's perspective. Journal of the Eastern Asia Society for Transportation Studies, 6, 4160-4175

Local Binary CNN for Diabetic Retinopathy Classification on Fundus Images

Peter Macsik, Jarmila Pavlovicova, Jozef Goga, Slavomir Kajan

Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava, Ilkovicova 3, 812 19 Bratislava, Slovakia
e-mail: peter.macsik@stuba.sk; jarmila.pavlovicova@stuba.sk; jozef.goga@stuba.sk; slavomir.kajan@stuba.sk

Abstract: Diabetic retinopathy (DR), is currently one of the major causes of preventable blindness, worldwide. With an early diagnosis and proper treatment of this eye disease, we can prevent the spread of diabetic retinopathy. In this paper, we propose a new alternative of local binary convolutional neural network (LBCNN) deterministic filter generation which can approximate the performance of the standard convolutional neural network (CNN) with less learnable parameters and also with less memory use, which can be helpful in systems with low-memory or low computational capacity, like smart-phones. We compare our scheme with standard CNN and LBCNN that uses stochastic filter generation strategy on retinal fundus image datasets in case of binary classification into healthy and damaged classes. These experiments are also evaluated according to the standard criteria used in medical applications, such as, overall accuracy, specificity, sensitivity and predictive values. On the small dataset (Aptos), one of our proposed LBCNN architectures outperformed all of the other deep learning models examined.

Keywords: CAD (Computer-aided diagnostics); Binary classification; Memory reduction; Learnable parameters

1 Introduction

Nowadays diabetes mellitus (DM) is a global disease [1] and the number of patients will probably increase in the future [2] [3]. On the other hand, diabetic retinopathy is the specific microvascular complication of DM and every third diabetic is affected by DR [1] and unfortunately is at risk of developing DR.

DR is an eye disease that can cause irreversible eye damages (i.e. blurred vision, black shapes, or dots in the vision area), in the worst cases even blindness, however, it could be preventable for in time diagnostic and proper treatment [4]. Diabetic retinopathy is classified according to symptoms and severity into two main groups, non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR). Subsequently, these stages are divided in more detail by individual symptoms according to the International Classification of Diabetic Retinopathy (ICDR) [1] scale. The international scale (ICDR) divides DR into five classes according to the severity of the disease. The definition of the ICDR scale is described in Table 1. Examples of fundus images from each ICDR class are shown in Fig. 1.

Table 1

Description of DR stages, ICDR scale [1]

Disease Severity Level	Findings Observable upon Dilated Ophthalmoscopy
No DR	No abnormalities
Mild NPDR	Microaneurysms only
Moderate NPDR	Microaneurysms and other signs, but less than severe NPDR
Severe NPDR	Moderate NPDR with any of the following: • Intraretinal hemorrhages (≥ 20 in each quadrant) • Definite venous beading (in 2 quadrants) • Intraretinal microvascular abnormalities (in 1 quadrant) • and no signs of proliferative retinopathy
Proliferative DR	Severe NPDR and 1 or more of the following: • Neovascularization • Vitreous/preretinal hemorrhage

Increased prevalence of DM leads to increased query for DR screening. Due to an increased number of patients with DM (respectively with DR) and an insufficient number of clinicians, there is more pressure on healthcare systems to find acceptable automatized DR screening methods with minimized costs.

In this paper, we propose the binary classification (especially classification into the healthy and damaged groups) with a memory-efficient CNN alternative which is called LBCNN and compare its performance with the standard CNN network. Memory-efficient CNN network alternatives are important for devices that are not equipped with sufficient memory. In our case, it could be for instance smart-phone which is one of the possible solutions how to avoid high costs and keep the comfort for the patient, but also enable regular DR screening. We tested our network on two fundus image databases (EyePACS [5] and Aptos [6]) of different sizes and showed computational efficiency of our solution in case of limited amount of training data.

The rest of this paper contains an overview of the related work in Section 2 and used datasets and image augmentations in Section 3. In Section 4 we introduce used methods for classification like standard CNN (ResNet18), LBCNN with stochastic filter generation mode and with proposed LBCNN with deterministic filter generation mode. In Section 5 there is the description of experiments and finally, we present our conclusions.

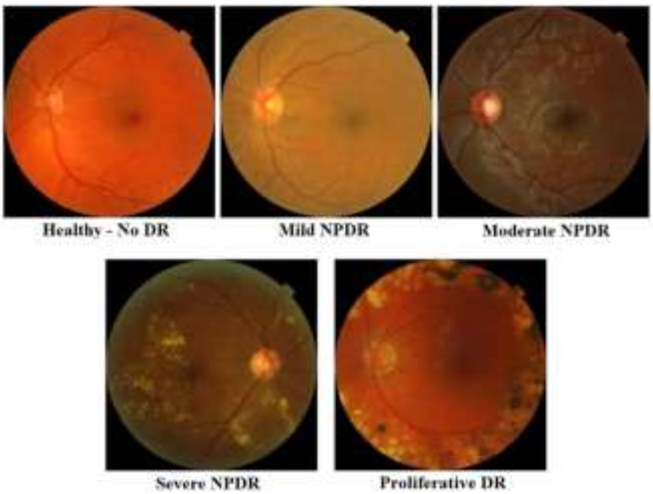


Figure 1
Examples of fundus images of different stages of DR

2 Related Works

The first attempts to automatically classify diabetic retinopathy were reported in the 1990s when Gardner and his colleagues described the usage of an artificial neural network, which was able to detect diabetic retinopathy with 88% sensitivity and 83% specificity compared to an ophthalmologist [7]. Certainly, since the first attempt, there were created many new applications for the classification of DR with different algorithms i.e. random forest classifier, support vector machine, or regression tree classifier reviewed in [8].

Nowadays, deep learning (DL) algorithms are the cost-effective solutions that could help to solve this problem. DL is a subarea in artificial intelligence (AI). On the other hand, CNN models belong to DL algorithms, that can be used among many others for image classification with repetitive analysis and compare the output with a standard (such as a human grader) and make self-correction in case of error. Several studies have confirmed successful results in the development of DL algorithms that have been able to identify DR without any need to have some specific properties of DR in advance. There are several ways to classify DR, e.g. by two (healthy, damaged), three (healthy, NPDR, PDR) or five (ICDR) classes [1].

For example, Islam et al. [9] developed two binary classification models. First, for detecting the presence of the disease (healthy vs damaged) and the second one for grading its severity (grades 0, 1 vs 2, 3, 4). In another study by Hagos et al. [10], the authors achieved 90.9% accuracy in binary classification with Inception-V3 [11] pre-trained and fine-tuned with a reduced dataset had 2500 fundus images.

Bodapati et al. [12] proposed a solution for binary classification (healthy versus diseased) and for the classification of the severity of DR (5 class - ICDR). They used different CNN architectures as feature extractor that were fused with deep neural network. Li et al. [13] proposed DR severity classification and an additional class (6 in total) to classify ungradable images as well. They trained many CNN architectures like VGG-16 [14], DenseNet-121 [15], GoogLeNet [16], ResNet-18 [17] where the best results in accuracy were achieved by ResNet-18 architecture.

On the other hand, one of the most challenging problems in designing robust DL methods, especially based on CNN models with deeper architectures, is the acquisition of huge volumes of labelled fundus images on pixel-level and with image-level annotations. The main issue is not the availability of huge datasets, but the annotation of these images, which is expensive and requires the services of expert ophthalmologists [18]. The solution could be a deep model, which is able to learn from limited data, and this is also an important area of research not only for the diagnosis of DR, but generally for medical image analysis, as well. To deal with this problem, we introduce a modified CNN model that has comparable performance with standard CNN but involves reduced number of learnable parameters.

Our research was inspired by the work of Juefei-Xu et al. [19]. They developed an efficient alternative to convolutional layers in standard CNN. This layer is called the local binary convolution (LBC) layer, which was motivated by local binary patterns (LBP) [20], a very efficient visual descriptor used for classification in computer vision. They called CNN with LBC layers LBCNN. In experiments, the LBCNN network was used for the classification task on ImageNet database [21]. The LBC layer comprises of fixed sparse pre-defined binary convolutional filters, which are fixed during the training process, a non-linear activation function and a set of learnable weights. The weights combine the activated filter responses to approximate the corresponding activated filter responses of a standard convolutional layer. The LBC layer affords significant savings, 9× to 169× in the number of learnable parameters compared to a standard convolutional layer (more details in Section 4.2). These parameter savings reduce memory and disk space requirements, which is beneficial for devices with lower computational power, e.g. smart phones [19]. Besides, smart phone DR screening is also a popular research field [22-24]. For example, Rajalakshmi et al. [24] assessed the role of AI based automated software for the detection of DR and sight-threatening DR fundus photography taken by a smartphone-based device and validated it against ophthalmologists grading.

For this reason, we applied the original LBCNN for DR classification and also we introduce an extension for the deterministic LBC layer with added Prewitt filters [25]. Thus, the original LBP filter base was extended for edge detection leading to improvement in feature extraction. The expected methodological scientific contributions of the paper were as follows:

- Experimental proof that original LBCNN with stochastic filters is usable also for binary DR classification and can achieve comparable results to standard CNN. For comparison, we use ResNet18 architecture as the best performing architecture in the study by Li et al. [13].
- Additional memory saving by application of a deterministic fixed filter base in LBCNN instead of stochastic filters while achieving comparable results to baseline LBCNN with stochastic filters. In this case it is not necessary to save the filter base for further reuse.

3 Datasets

We propose image classification of fundus images obtained by a fundus camera. Fundus images show the interior surface of the eye, opposite to the lens. In our work, we chose EyePACS [5] and Aptos [6] from freely available fundus eye databases. These two databases differ in size markedly. Difference in size allows us to demonstrate the benefits and the drawbacks of proposed methods compared to standard CNN.

3.1 EyePACS

EyePACS database [5] contains color fundus images which were divided by ophthalmologists into five classes (ICDR) according to the grade of DR retinal damage. EyePACS provided this database in 2015, for Kaggle [26] competitors. The aim of this competition was to design the best possible automated detection system of DR symptoms [1]. Original database contains 35126 training images with different image resolutions with following grade distribution: No DR 25810 (grade 0), Mild NPDR 2 443 (grade 1), Moderate NPDR 5 292 (grade 2), Severe NPDR 873 (grade 3), Proliferative DR 708 (grade 4). Due to few images in classes 3 and 4 which indicate unbalance between classes, we augmented (similarly like in [27]) this part of the dataset by adding images from the EyePACS testing dataset [5]. In case of class 3 it was 2087 images and in case of class 4 it was 1914 images. Since the dataset contains also left and right eyes, we used mirroring to double the number of images. After this augmentation, we observed 25790 images of healthy and 23472 images with DR symptoms.

3.2 Aptos

Asia Pacific Tele-Ophthalmology Society 2019 Blindness Detection Dataset [6] was divided as well as EyePACS dataset into five classes. Public Aptos database contains 3662 images with various resolutions (up to 3216×2136) with following DR grade distribution: No DR 1805 (grade 0), Mild NPDR 370 (grade 1), Moderate NPDR 999 (grade 2), Severe NPDR 193 (grade 3), Proliferative DR 295 (grade 4).

In the context of our binary classification, after merging retinal images with DR, we obtained 1805 images of healthy retina and 1857 of damaged retinal images with signs of DR.

3.3 Fundus Image Preprocessing

In both databases, it is possible to observe black borders around the eye fundus. However, from the point of view of CNN network training, these black borders do not contain any important information, so in order to reduce the size of input images and at the same time reduce computational complexity, it is appropriate to trim them. For this reason, before the training process we cropped black borders automatically with an adaptive method, similarly as in the study by Shao et al. [28]. After border cropping, images were resized to 300×300 pixels to reach uniform image resolution. An example of the cropped and resized image is shown in Fig. 2.

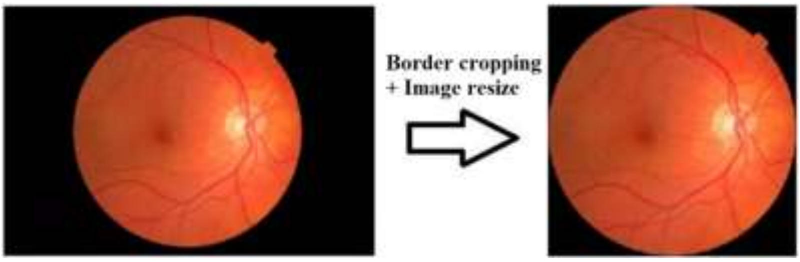


Figure 2
Fundus image preprocessing example

4 Methods

As a classification algorithm baseline model, we used well known CNN network and its new alternative LBCNN [19] in frame of architecture ResNet18 (Fig. 3). LBCNN was born from the idea of combination LBP descriptor and CNN architecture. The main advantage of LBCNN is the potential to achieve comparable results with CNN architecture with the benefits of less learnable parameters and lower memory requirements.

4.1 CNN

Convolutional neural networks are widely used deep learning models, which achieve high popularity for image classification tasks. They are mostly based on computational layers like convolutional or pooling layer and activation functions like ReLU or sigmoid. These networks have randomly initialized convolutional filters which are optimized during the training for feature extraction. One of the first

CNN was developed by Yann LeCun et al. which was called LeNet [29]. During the years many architectures of CNNs were published, i.e. AlexNet [30], ResNet [17], VGG [14], etc.

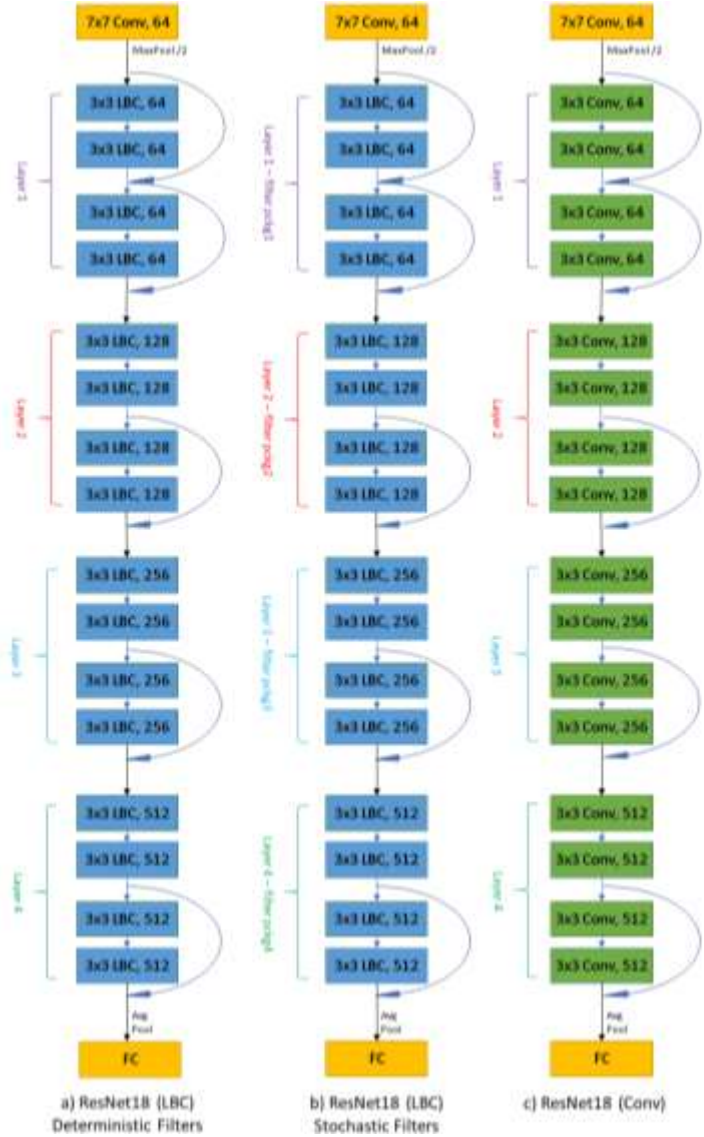


Figure 3
Visualization of used ResNet18 models: a) LBC layered with deterministic filters; b) LBC layered with stochastic filters; c) with standard convolutional layers

These architectures have different numbers and types of layers with different connections. In most cases, these architectures have a huge number of learnable parameters and certainly, this could lead to increased memory requirements. In this paper we compared one alternative of CNN with reduced number of learnable parameters. This network is called LBCNN [19]. As an experimental CNN architecture, for results comparison, we chose ResNet18 [17] implemented in PyTorch framework [31] (model c) in Fig. 3).

4.2 Local Binary CNN - LBCNN

LBCNN [19] is an alternative to the standard CNN which can approximate the performance of CNN with less learnable parameters. It was born from the idea of LBP convolution which has 8 special binary non-learnable filters, activation function, and binary weights for a linear combination. These factors were generalized to the m binary fixed filters (they are not learnable). The linear combination part of the layer was generalized from binary numbers to the real values. This linear combination with real values was applied as pointwise convolution (convolution with 1×1 sized filters) and this is the only part where this layer can learn [20]. Such layer is called the LBC layer and CNN with these LBC layers is called LBCNN. LBC layer function can be expressed by the following equation:

$$x_{l+1}^{t} = \sum_{i=1}^{m} \sigma\left(\sum_{s} b_i^s * x_l^s\right) \cdot V_{l,i}^{t} \tag{1}$$

where t and s represent the number of input and output channels, m is the number of fixed filters (b_i, $i \in [m]$), x_l is the input from l^{th} layer and x_{l+1} is the output from layer, and consequently it is the input into layer $l + 1$. $V_{l,i}$ are weights in pointwise convolution. The activation function is σ (we used ReLU). Operator * stands for standard 2D convolution and operator · denotes pointwise convolution. Visualization of a single LBC layer is shown in Fig. 4.

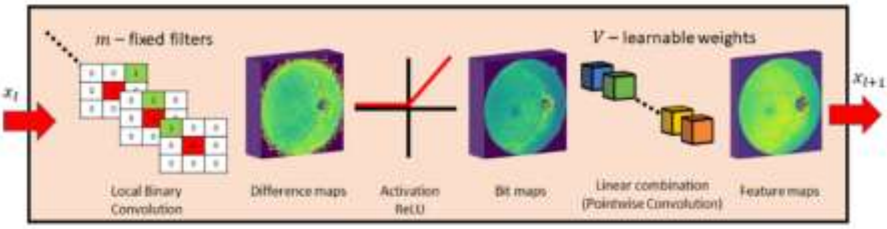


Figure 4

Single LBC layer (m - number of fixed filters, V - weights for linear combination)

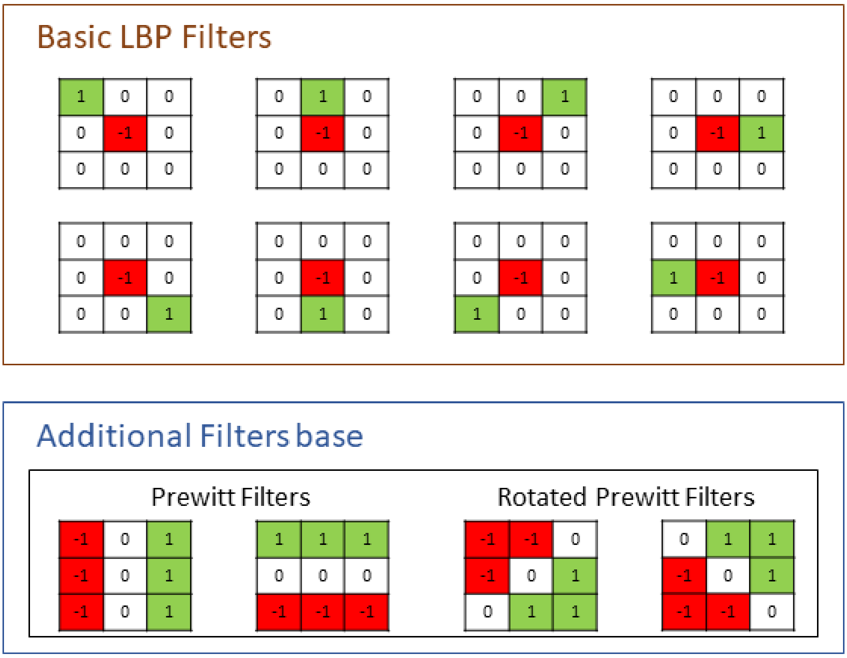
LBCNN saves parameters through binary non-learnable filters which can be generated in two ways, one is deterministic and the second one is stochastic. In this paper, we used a deterministic and also stochastic filter generation strategy.

4.2.1 Stochastic LBC Filters

We used stochastic fixed filter generation described in [19]. Memory savings were achieved here by the ability to share fixed filters across layers with the same dimensions. We generated a new package of fixed filters for every layer (model b) in Fig. 3) and we shared them between LBC layers. First, filter generation sparsity must be defined, which represents the ratio between zero values and non-zero values in the filter. If the value is non-zero it has value 1 or -1, according to the Bernoulli distribution. We used stochastic filters with a sparsity of 0.5, that was determined in original paper [19] as a good standard value.

4.2.2 Deterministic LBC Filters

Deterministic filter generation strategy can save extra memory compared to stochastic filter generation. In case of deterministic filters, it is not necessary to save fixed filters after training because we know how they look like. In case of stochastic generation, it is necessary to save all fixed filters for further model re-use due to random factor. In this paper, we used original LBP filters with some additional deterministic fixed filters in order to increase the filter base. Additional filters are shown in Fig. 5.



Basic LBP Filters

1	0	0
0	-1	0
0	0	0

0	1	0
0	-1	0
0	0	0

0	0	1
0	-1	0
0	0	0

0	0	0
0	-1	1
0	0	0

0	0	0
0	-1	0
0	0	1

0	0	0
0	-1	0
0	1	0

0	0	0
0	-1	0
1	0	0

0	0	0
1	-1	0
0	0	0

Additional Filters base

Prewitt Filters

-1	0	1
-1	0	1
-1	0	1

1	1	1
0	0	0
-1	-1	-1

Rotated Prewitt Filters

-1	-1	0
-1	0	1
0	1	1

0	1	1
-1	0	1
-1	-1	0

Figure 5

Base filters used in LBC layers. Basic 8 LBP filters extended by another 4 filters, Prewitt and rotated Prewitt filters

This additional base contains Prewitt filters [25] and rotated Prewitt filters for edge recognition which can be important near vessels and borders of the eye and also for better detection of circle-shaped disease signs i.e. microaneurysms. We used LBCNN with deterministic filters with 2 setups. In the first, we used 12 filters, as in Fig. 5. In the second, we doubled the number of fixed filters to 24, by keeping the original 12 and the expansion was done by swapping values -1 and 1 in every filter.

4.2.3 Number of Learnable Parameters

The major advantage of LBC layers application is the reduction of learnable parameters number by the preservation of similar learning ability. If we assume that convolutional filters do not have bias terms, comparison between learnable parameters in each LBC and Conv layer can be expressed with the following equation:

$$\frac{CNN\ params.}{LBCNN\ params.} = \frac{p \times h \times w \times q}{m \times q} \qquad (2)$$

where m in our case was 12, 24, and 72 which is the number of fixed filters in LBC layers. Next parameters h and w stand for the height and width of fixed LBC filters (both are 3 in our case), respectively. Parameters p and q stand for the number of input and output channels. The accurate parameter difference for the models is shown in Table 2. This table contains model name and number of learnable parameters (Params).

Table 2

Learnable parameters. [number]f - number of filters, sto. - stochastic, det. - deterministic)

Model	Params [million]		
Standard CNN – ResNet18	11.178		
	12f	24f	72f
LBCNN – ResNet18 sto.	–	0.288	0.472
LBCNN – ResNet18 det.	0.242	0.288	–

4.2.4 Memory Size Difference

Deterministic filter generation strategy has the advantage in memory saving compared to the stochastic generation as we described above, as there is not necessary to store them. It means that alternatively, we could generate them programmatically in a predefined order. It is sufficient to save only learnable parts of the model, which are pointwise convolutional layers. However, in case of stochastic filter generation, it is important to save fixed filters, otherwise, further model re-use is impossible. To compare the memory requirement of these models in the PyTorch framework [15] we saved networks in the *.ckp file format, where we can save trained parameters and fixed filters for further re-use. This comparison is shown in Table 3. It contains the memory size requirements for standard CNN ResNet18 and each setup of LBCNN.

Note: The specific memory size may vary from different hardware and software factors.

Table 3
Memory size difference. ([number]f-number of filters, sto. - stochastic, det. - deterministic)

Model	Size [KB]		
Standard CNN – ResNet18	131106		
	12f	24f	72f
LBCNN – ResNet18 sto.	–	4325	8105
LBCNN – ResNet18 det.	2972	3512	–

5 Experiments and Results

In this chapter, we present our experiments of two approaches with proposed architectures. In the first case we did experiments on basic single model classification which is favorable in case of low memory and computational capacity. In the second approach, we experiment with an ensemble of more models that can offer improvements in classification accuracy with minimal increase in the number of parameters and memory requirements.

5.1 Single Model Classification

Firstly, we have made hyperparameter tuning with grid search method and empirically discovered the best training setup for selected models. We have tested different weight optimization methods (such as Adagrad, Adadelta, Adam, AdamW, and Nadam [32-35]) with different hyperparameters. We achieved the best results with Nadam optimizer, with learning rate 0.001 and with fixed number of epochs, 30 and 40 for EyePACS and Aptos, respectively. Other parameters were kept default as in PyTorch implementation of Nadam, which is on GitHub repository [36]. This hyperparameter tuning was made on CNN (ResNet18), and for objective comparison of models performance, we kept this setup for LBCNN models too. We made experiments on all the above-mentioned datasets, where we divided datasets into 80-20% ratio for the training and testing with random data selection in all cases. For every model, we made 30 repeated runs with the setups described above. After 30 runs we chose 10 best models based on the test accuracy, and evaluated the obtained results.

Results of our experiments for the smaller dataset (Aptos) are shown in Table 4, and for the bigger dataset (EyePACS) in Table 5. Tables contain evaluation metrics used in medicine like accuracy, sensitivity (sens), specificity (spec), negative predictive value (NPV), and positive predictive value (PPV). These metrics were calculated as an average of 10 best models. The above-described metrics for a single model were calculated from confusion matrices through the use of TP, FP, TN, FN values (F-False, T-True, P-Positive, N-Negative).

Specifically, in medical classification tasks, confusion matrix values can be described as follows:

- True Positive (TP): Damaged image correctly identified as damaged
- False Positive (FP): Healthy image incorrectly identified as damaged
- True Negative (TN): Healthy image correctly identified as healthy
- False Negative (FN): Damaged image incorrectly identified as healthy

These metrics are expressed by the following equations:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)$$

$$spec = \frac{TN}{TN+FP} * 100 \quad (4)$$

$$sens = \frac{TP}{TP+FN} * 100 \quad (5)$$

$$NPV = \frac{TN}{TN+FN} * 100 \quad (6)$$

$$PPV = \frac{TP}{TP+FP} * 100 \quad (7)$$

To express the relation between specificity and sensitivity we also used an evaluation metric called AUC (Area under receiver operating characteristic curve) [37]. AUC is included in Tables 4 and 5 also as median accuracy and standard deviation (std) of the 10 best models. Tables also contains best and median accuracy. Figs. 6-7 show boxplot visualization of the 10 best models performance.

Table 4

Results of 10 best experiments on Aptos dataset (det. - deterministic, sto. - stochastic, f – filters, acc. - accuracy)

	Models				
	CNN ResNet18 [17]	**LBCNN (sto. 24f) [19]**	**LBCNN (sto. 72f) [19]**	**LBCNN (det. 12f) [ours]**	**LBCNN (det. 24f) [ours]**
mean acc. [%]	95.59	95.57	95.44	95.89	**96.58**
std	0.556	0.4616	**0.362**	0.3832	0.4022
median acc. [%]	95.29	95.36	95.36	95.91	**96.52**
best acc. [%]	96.73	96.32	96.04	96.45	**97.41**
auc	0.979	0.9803	0.979	0.9832	**0.9871**
spec [%]	95.94	96.01	95.74	95.96	**96.59**
sens [%]	93.77	93.67	93.22	94.03	**94.63**
NPV [%]	93.95	93.56	93.39	94.05	**94.35**
PPV [%]	95.71	95.83	95.75	96.08	**96.73**

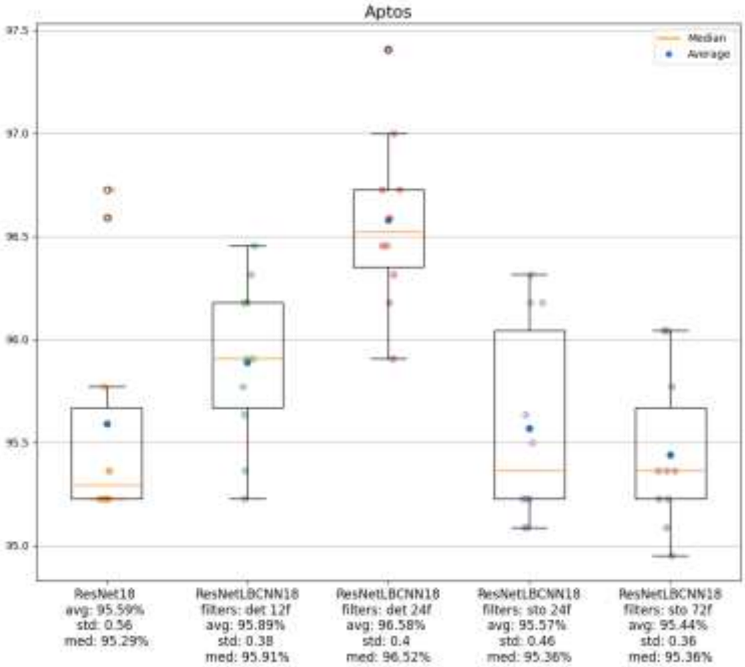


Figure 6

Boxplot visualization of reached results on Aptos for used models (best 10 experiments)

Table 5

Results of 10 best experiments on EyePACS dataset (det. - deterministic, sto. - stochastic, f - filters, acc. - accuracy)

	Models				
	CNN ResNet18 [17]	**LBCNN (sto. 24f) [19]**	**LBCNN (sto. 72f) [19]**	**LBCNN (det. 12f) [ours]**	**LBCNN (det. 24f) [ours]**
mean acc. [%]	**91.12**	90.22	90.4	88.73	89.71
std	**0.1828**	0.2484	0.2519	0.3696	0.2475
median acc. [%]	**91.11**	90.14	90.33	88.66	89.65
best acc. [%]	**91.44**	90.73	91.03	89.34	90.31
auc	**0.9725**	0.9698	0.9706	0.9605	0.9661
spec [%]	93.32	93.06	**93.41**	92.23	93.08
sens [%]	**87.82**	86.27	86.72	84.56	85.78
NPV [%]	**89.34**	88.14	88.5	86.68	87.65
PPV [%]	92.22	92.05	**92.28**	90.79	91.91

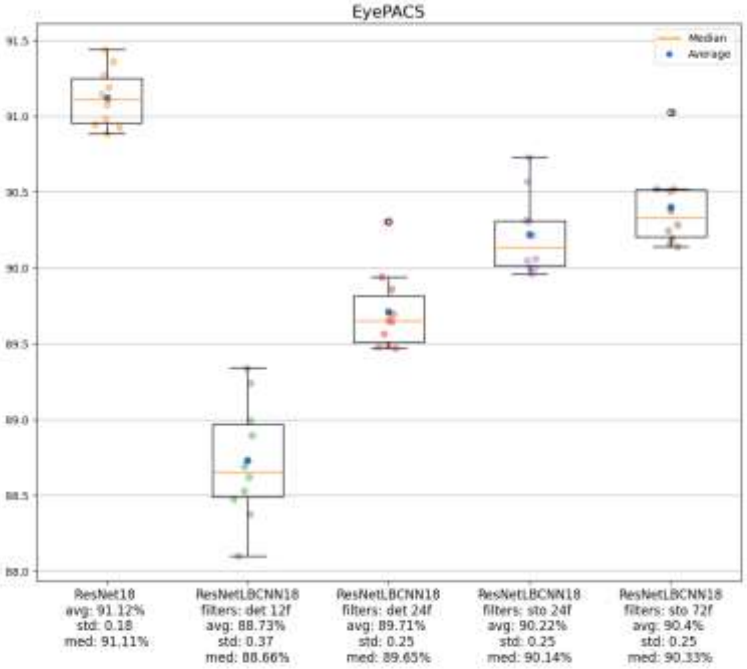


Figure 7

Boxplot visualization of reached results on EyePACS for used models (best 10 experiments)

5.2 Ensemble Classification

A different, interesting practical approach, could be the ensemble creation of standard CNN and LBCNNs, however, this ensemble model [38] will require a larger model size, over the single standard CNN, but on the other hand, we could achieve classification improvement, using minimal parameters and minimal model size increases. As a demonstration, we used ensemble of 3 models (LBCNN deterministic with 24 filters, LBCNN stochastic with 24 filters and standard ResNet18) (Fig. 8). We used the Model Averaging Ensemble, to combine particular predictions, which means, every single model has an equal impact (weight), on the final prediction.

We made these experiments on EyePACS database where ensemble of 3 models with equal weights in ensemble produced the following results: mean and median of the best 10 experiments were 92.00% and 91.97%, respectively. The best result we achieved was 92.39%. It means +0.88%, +0.86%, +0.95% improvement of accuracy on average, median, and on the best model with adding just 7 837 KB of memory or in learnable parameters, it means +0.488 million additional learnable parameters.

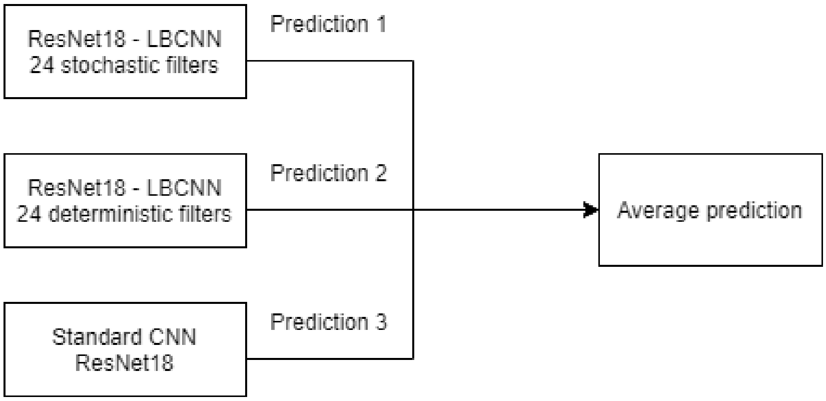


Figure 8
Visualization of Averaging Ensemble Model

Conclusions

In our experiments, we proposed the additional deterministic filters application, in LBCNN, to achieve a more accurate DR image classification, for healthy or damaged classes. Specifically, this means that we extended a filter base of 8 LBP filters by 4 Prewitt filters and then we doubled the number of filters by swapping non-zero values in each filter. Thus, we effectively created 24 fixed filters. This deterministic filter generation can decrease the parameters memory requirement, compared to stochastic filters, because there is no need for fixed filter storing for further trained model re-use. This approach can be useful for low-memory devices, such as, smart-phones. It can be a very cost-effective solution of regular DR screening. However, in general, the selection of the deterministic filters base can also be a weakness, for example, in the case that the selected filters are not optimized for the given classification task, the performance of this LBCNN can be even worse, compared to the standard CNN.

Based on our experiments on fundus image datasets, we can establish that LBCNN, with both strategies of filter generation (stochastic and deterministic), can approximate the performance of a standard CNN network for binary DR classification, moreover, it saves a significant amount of learnable parameters and decreases memory requirements. Specifically, in experiments on the smaller dataset (Aptos), performance of the LBCNN, with 24 deterministic filters, gave the best results, except for the standard deviation, where LBCNN with 72 stochastic filters achieved the best results, however, the difference was negligibly low. In case of the larger dataset (EyePACS), performance of the LBCNN with 24 deterministic filters, was slightly worse, because of an insufficient number of parameters, but it can be said that there is still a good trade-off between performance and the memory requirements. In this case, the standard CNN achieved the best results. This indicates that using LBCNN with deterministic filters is fully applicable for classification tasks, where only small datasets are available.

LBCNN with deterministically or stochastically generated filters are also usable separately and also as a part of an ensemble model, as a mechanism of improvement. Certainly, this option requires more memory compared to the single CNN, but if memory allows for it, it can be an alternative to improve classification accuracy with a minimal memory increase. This improvement was also demonstrated in our experiments, where we combined 3 models (LBCNN with 24 stochastic filters, LBCNN with 24 deterministic filters and a standard CNN-ResNet18). Conversely, using the dataset EyePACS, we achieved almost +1% improvement in accuracy, compared to the best classifier of the group.

As future work, we plan to experiment with a greater ensemble of models, with the purpose to find optimal weights for classification performance, with minimum memory requirement target.

References

[1] International Council of Opthalmology, ICO Guidelines for Diabetic Eye Care, (2017) 40

[2] D. R. Owens, J. Dolben, S. Young, R. E. J. Ryder, I. R. Jones, J. Vora, D. Jones, D. Morsman, T. M. Hayes, Screening for Diabetic Retinopathy, Diabetic Medicine. 8 (1991) S4–S10. https://doi.org/10.1111/j.1464-5491.1991.tb02148.x

[3] D. C. Klonoff, The Increasing Incidence of Diabetes in the 21st Century, Journal of Diabetes Science and Technology. 3 (2009) 1-2. https://doi.org/10.1177/193229680900300101

[4] P. Vashist, S. Singh, N. Gupta, R. Saxena, Role of Early Screening for Diabetic Retinopathy in Patients with Diabetes Mellitus: An Overview, Indian Journal of Community Medicine : Official Publication of Indian Association of Preventive & Social Medicine. 36 (2011) 247-252. https://doi.org/10.4103/0970-0218.91324

[5] EyePACS, Diabetic Retinopathy Detection - Kaggle, 2015. https://www.kaggle.com/c/diabetic-retinopathy-detection (accessed January 27, 2021)

[6] APTOS, APTOS 2019 blindness detection, 2019. https://www.kaggle.com/c/aptos2019-blindness-detection

[7] G. G. Gardner, D. Keating, T. H. Williamson, A. T. Elliott, Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool., British Journal of Ophthalmology. 80 (1996) 940–944. https://doi.org/10.1136/bjo.80.11.940

[8] S. Vadloori, Y.-P. Huang, W.-C. Wu, Comparison of Various Data Mining Classification Techniques in the Diagnosis of Diabetic Retinopathy, Acta Polytechnica Hungarica. 16 (2019) 27-46

[9] S. M. S. Islam, M. M. Hasan, S. Abdullah, Deep Learning based Early Detection and Grading of Diabetic Retinopathy Using Retinal Fundus Images, ArXiv:1812.10595 [Cs] (2018) http://arxiv.org/abs/1812.10595 (accessed December 27, 2021)

[10] M. T. Hagos, S. Kant, Transfer Learning based Detection of Diabetic Retinopathy from Small Dataset, ArXiv:1905.07203 [Cs] (2019) http://arxiv.org/abs/1905.07203 (accessed December 8, 2021)

[11] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the Inception Architecture for Computer Vision, ArXiv:1512.00567 [Cs] (2015) http://arxiv.org/abs/1512.00567 (accessed February 12, 2022)

[12] J. D. Bodapati, V. Naralasetti, S. N. Shareef, S. Hakak, M. Bilal, P. K. R. Maddikunta, O. Jo, Blended Multi-Modal Deep ConvNet Features for Diabetic Retinopathy Severity Prediction, Electronics. 9 (2020) 914. https://doi.org/10.3390/electronics9060914

[13] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu, H. Kang, Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening, Information Sciences. 501 (2019) 511-522, https://doi.org/10.1016/j.ins.2019.06.011

[14] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, ArXiv:1409.1556 [Cs] (2015) http://arxiv.org/abs/1409.1556 (accessed February 4, 2021)

[15] G. Huang, Z. Liu, L. van der Maaten, K. Q. Weinberger, Densely Connected Convolutional Networks, ArXiv:1608.06993 [Cs] (2018) http://arxiv.org/abs/1608.06993 (accessed December 5, 2021)

[16] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going Deeper with Convolutions, ArXiv:1409.4842 [Cs] (2014) http://arxiv.org/abs/1409.4842 (accessed November 29, 2021)

[17] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016. https://doi.org/10.1109/CVPR.2016.90

[18] N. Asiri, M. Hussain, F. Al Adel, N. Alzaidi, Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey, Artificial Intelligence in Medicine. 99 (2019) 101701. https://doi.org/10.1016/j.artmed.2019.07.009

[19] F. Juefei-Xu, V. N. Boddeti, M. Savvides, Local binary convolutional neural networks, in: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017. https://doi.org/10.1109/CVPR.2017.456

[20] F. Juefei-Xu, M. Savvides, Weight-Optimal Local Binary Patterns, in: L. Agapito, M.M. Bronstein, C. Rother (Eds.), Computer Vision - ECCV 2014 Workshops, Springer International Publishing, Cham, 2015: pp. 148-159. https://doi.org/10.1007/978-3-319-16181-5_11

[21] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, ImageNet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009: pp. 248-255. https://doi.org/10.1109/CVPR.2009.5206848

[22] M. W. M. Wintergerst, D. K. Mishra, L. Hartmann, P. Shah, V. K. Konana, P. Sagar, M. Berger, K. Murali, F. G. Holz, M. P. Shanmugam, R. P. Finger, Diabetic Retinopathy Screening Using Smartphone-Based Fundus Imaging in India, Ophthalmology. 127 (2020) 1529-1538. https://doi.org/10.1016/j.ophtha.2020.05.025

[23] R. Rajalakshmi, S. Arulmalar, M. Usha, V. Prathiba, K. S. Kareemuddin, R. M. Anjana, V. Mohan, Validation of Smartphone Based Retinal Photography for Diabetic Retinopathy Screening, PLOS ONE. 10 (2015) e0138285. https://doi.org/10.1371/journal.pone.0138285

[24] Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence | Eye, (n.d.). https://www.nature.com/articles/s41433-018-0064-9 (accessed January 25, 2022)

[25] J. M. S. Prewit, Object enhancement and extraction, Picture processing and Psychopictorics, 1970

[26] Kaggle Competitions, (n.d.) https://www.kaggle.com/competitions (accessed February 12, 2022)

[27] S. Kajan, J. Goga, K. Lacko, J. Pavlovičová, Detection of Diabetic Retinopathy Using Pretrained Deep Neural Networks, in: Conference Cybernetics & Informatics, Velké Karlovice, Czech Republic, 2020

[28] F. Shao, Y. Yang, J. Qiuping, G. Jiang, Y.-S. Ho, Automated Quality Assessment of Fundus Images via Analysis of Illumination, Naturalness and Structure, IEEE Access. PP (2017) 1-1. https://doi.org/10.1109/ACCESS.2017.2776126

[29] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE. 86 (1998) 2278-2324. https://doi.org/10.1109/5.726791

[30] A. Krizhevsky, One weird trick for parallelizing convolutional neural networks, ArXiv:1404.5997 [Cs]. (2014) http://arxiv.org/abs/1404.5997 (accessed February 4, 2021)

[31] PyTorch, From research to production, (n.d.) https://pytorch.org/

[32] J. Duchi, E. Hazan, Y. Singer, Adaptive Subgradient Methods for Online Learning and Stochastic Optimization, Journal of Machine Learning Research. 12 (2011) 2121-2159

[33] M. D. Zeiler, ADADELTA: An Adaptive Learning Rate Method, ArXiv:1212.5701 [Cs] (2012) http://arxiv.org/abs/1212.5701 (accessed February 9, 2021)

[34] I. Loshchilov, F. Hutter, Decoupled Weight Decay Regularization, ArXiv:1711.05101 [Cs, Math] (2019) http://arxiv.org/abs/1711.05101 (accessed February 9, 2021)

[35] T. Dozat, Incorporating nesterov momentum into adam (2016)

[36] NAdam — PyTorch 1.10.0 documentation, (n.d.) https://pytorch.org/docs/stable/generated/torch.optim.NAdam.html#torch.optim.NAdam (accessed November 16, 2021)

[37] T. Fawcett, An introduction to ROC analysis, Pattern Recognition Letters. 27 (2006) 861-874. https://doi.org/10.1016/j.patrec.2005.10.010

[38] O. Sagi, L. Rokach, Ensemble learning: A survey, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 8 (2018) e1249

Performance Improvement of Face Recognition Method and Application for the COVID-19 Pandemic

Suparna Biswas

Faculty of Electronics and Communication Engineering department at Gurunanak Institute of Technology, 157/F, Nilgunj Rd, Sahid Colony, Panihati, Sodepur, Kolkata, West Bengal, 700114, India, suparna.biswas@gnit.ac.in

Abstract: In this paper, a novel framework is introduced by combining compressive sensing(CS) theory, digital curvelet transform, and Principal Component Analysis to improve the performance of face recognition method. CS is a highly attractive approach in the field of signal processing, which provides an efficient way of data sampling at a lower rate than the Nyquist sampling rate. CS offers numerous advantages, like less memory storage, less power consumption and higher data transmission rate etc. Here, CS is used on the face images, which offers reduction in storage space and computational time. The use of curvelet transform provides dual benefits: (i) sparse representation (ii) improvement on detailed content. To extract the feature vector, the Principal Component Analysis is then applied. The Performance of the proposed face recognition method is computed by applying cross-validation technique, compressive sensing based classifier, neural network, Naive Bayes and Support Vector Machine classifier. The proposed technique can efficiently perform the face recognition, at a low computational cost. Extensive experiments, on ORL and AR face databases, are conducted to validate our claim. The proposed technique also recognizes face images more efficiently than the traditional PCA, with a 1.5% higher recognition rate, if a person wears a face mask, as protection from COVID-19

Keywords: Face Recognition; Compressive Sensing; COVID-19; Curvelet Transform

1 Introduction

Face recognition (FR) [1] is a widely used biometric based technique for identification of individuals in various places for security issues. FR offers various advantages over the other biometric based techniques (iris, retina, fingerprint, etc.) like face images can be captured with a low cost camera and no need for direct contact of the acquisition device. But it has some disadvantages too. One of the notable disadvantages of the FR system is that it is less reliable than other biometric based systems. Actually, different factors like quality of image or video, expression variation, Occlusion, illumination differences affect the recognition performance.

Identification through FR [2] [3] [4], is basically a matching problem, in which, test image is compared with the stored database. But sometimes automatic FR is very much challenging or hard task due to the variations of different factors like pose variations, expression variations, illumination differences, presence of noise, occlusion and blurring etc. Some form of pre-processing steps become required to reduce the noise, effect of variations in pose and illumination, which have impact on the choice of the recognition scheme. Automatic FR methods involve two important steps:

i) Extraction of features from face images

ii) Classifier design

However, classification result mostly depends on the feature extraction techniques. In traditional method of FR, pixels intensity is used as input features. However, the techniques are time consuming due to high dimensionality of input feature vector.

Various methods have been developed for extracting features in low dimensional space, such as Principal Component Analysis (PCA) [5], Linear Discriminant Analysis (LDA) [6] and Independent Component Analysis (ICA) [7]. PCA is most widely used method for feature selection and dimensionality reduction. However, PCA can’t handle variation in illumination and facial expression. LDA is another powerful tool for dimensionality reduction, widely used in FR. In case of traditional LDA, classification accuracy may degrade with the sample size. PCA can give outstanding performance than LDA, if the training dataset is small. Recently lots of FR methods have been developed applying the unsupervised statistical techniques. In unsupervised statistical technique, a set of basis images are used to represent faces as a linear combination of the basis images. Higher order statistics among the pixel values provide better basis images, which contain more important information for FR. ICA is one such method depicted as generalization of PCA and superior than PCA in case of illumination and expression variations of face images.

Wavelets [8] and the various variants, namely contourlet [9] curvelet [10] etc. are found to be efficient to analyze the high dimensional signals. Wavelets offer the benefits of multiscale analysis and time frequency localization of 2D image matrix. However, wavelets are ineffective while dealing with smooth contours in different directions unlike contourlet which can handle this issue due to additional properties of directionality and anisotropy. Moreover, wavelet transform can detect only point singularities but fails to detect curved singularities. To overcome the limitations of wavelet and contourlet, Candes and Donoho [10] presented curvelet transform (CVT) which has better capability to represent edges and other singularities along curves. CVT represents the line, the edges and the curvature precisely through compact representation using less number of coefficients.

During the recent years, CS theory has received remarkable attention in the research area of FR [1]. CS theory reinvents FR technique applying sparse representation theory. Usually, in the case of FR technique using CS theory, the query image is presented as a sparse linear combination of training images. This method showed robustness in presence of noise. In CS, sparsification is an important step and degree of sparsity has a significant role in reconstruction process. So investigation of suitable sparse representation is the most vital task for FR technique, based on CS.

The global problem at present is COVID-19 caused by corona-virus which led to the worldwide lockdown. The common people are the worst sufferer having no work. Though intensive research and development of vaccines is currently underway in Russia, UK, USA and other countries giving the common people a ray of hope, yet they have decided to return to their work to earn their livelihood. Considering the present scenario wearing mask and protective gear has become mandatory for all but it may create some security issues as it is hiding the face of individuals.

The objective of the present work is to improve feature information of face images in order to recognize the individuals efficiently. In this paper features of CS, CVT and PCA are exploited to develop a new FR technique. Here, CVT has been used to perform dual role, first one is sparse representation and another is enhancement of CS reconstructed face images. To extract the features in low dimensional feature space PCA has been used on the enhanced image. Comprehensive simulations are conducted on two online accessible datasets, applying different classification technique to demonstrate the supremacy of our proposed scheme. Our proposed technique also tested on face with mask to combat with corona virus as a protective measure.

The remaining part of the paper is structured as follows: Section 2 provides a literature review on FR method. In Section 3 the proposed FR method is described clearly and Section 4 provides the experimental results and discussion. In Section 5 the modified proposed method is discussed to combat with COVID-19. Finally, the paper is concluded in Section 6.

2 Scientific Literature Review

A brief review of FR and its superiority followed by the benefaction of the proposed method has been discussed in this section.

2.1 Related Work

Recognition of human faces utilizing PCA was first proposed by Sirovich and Kirby [11] in 1987. Some recent advancement on PCA based algorithms includes symmetrical PCA [12], two-dimensional PCA (2D-PCA) [13-14], weighted modular PCA [15], Kernel PCA [16] and diagonal PCA [17]. The method 2D-PCA is dependent on 2D image matrix. Hence, it does not need to transform the 2D image matrix into a vector, prior to feature extraction. Here image covariance matrix directly constructed from the 2D matrix and this technique is computationally more adequate than traditional PCA. In [17], fit has been reported that diagonal PCA (DiaPCA) is more accurate than both PCA and 2D-PCA. Improved FR performance was observed by combining the DiaPCA and 2D-PCA. Recently published other performance improvement techniques are [18-19]. In [18] a Local Binary Pattern (LBP) Histogram based technique and in [19] LBP and Support Vector Machine (SVM) has been used to improve the recognition rate.

Sparse representation based classification (SRC) technique for FR was introduced by Wright et al. [1]. In [1] query image is represented as sparse linear combination of the training images and applying l1 - minimization technique [20], the sparsest coding vector has been achieved. Then the classification of test image was based on minimum representation error. Wright et al. [1] declared that for large dimension of feature vector SRC is independent of feature types. To overcome the difficulty of occlusion and corruption Wright et al. [1] introduced an occlusion dictionary and showed that this technique is robust for small variations of pose and displacement.

Yang et al. [20] proposed a FR technique utilizing l1 norm minimization SRC algorithm. In this paper Gabor features are extricated from the local regions of face images, which are slightly sensitive to variations of pose, illumination and expression than the holistic features. This FR method [20] based on Gabor feature has improved the classification rate over the conventional SRC based technique and increased the computational speed in presence of occlusion.

Yang et al. [21] implemented a robust sparse coding (RSC) method to recognize the face images robustly. A suitable weight function is designed for RSC, to obtain better performance than the existing SRC [20] based method, with the intricate variations of faces. Assuming that the noise term has a sparse representation, a correntropy based sparse representation (CESR) technique was proposed in [22]. CESR technique can efficiently handle the non-Gaussian noise and yields better results for the scarf occlusion problem in FR.

Huang et al. [23] introduced a FR method which is transformation-invariant SRC technique. Zhou et al. [24] integrated Markov random model with SRC technique to recognize the face images under contiguous occlusion. Wagner et al. [25] handled pose and illumination variation FR problem by introducing SRC framework. In [26] a discriminative dictionary learning technique was proposed to increase the efficiency of FR. Yi-Haur et al. [27] developed a FR technique named as maximum probability of partial ranking on the framework of SRC. In [27] 2D-PCA and 2D-LDA feature extraction techniques has been applied on two face databases ORL and Yale B. This technique [27] showed significant improvement in recognition accuracy greater than the traditional PCA, LDA, two dimensional PCA and LDA based techniques. Considering the correlation and sparsity Wang et al. [28] presented a FR technique which is based on adaptive sparse linear model. It has been noticed that this sparse linear model behaves like SRC if the training samples are almost uncorrelated. But if the training samples are highly correlated then [28] this technique behaves as collaborative representation based classification (CRC). This technique has showed better performance over the Neural Network (NN), SRC and CRC techniques.

Some of the recent works on FR are [29] [30] [31] [32] [33]. In [29], the author presented a random sampling patch-based FR technique to cope up with the problem of occlusion. In the same year Wang et al. [30] presented another sparse representation based FR technique to overcome the same problem. Recently deep learning is widely used in different pattern recognition problems, due to its high recognition accuracy. Following this trend, Feng et al. [31] proposed a deep learning based Robust LSTM autoencoder to handle the occlusion. To address the issues of pose variation, Kishor et al. [32], presented a FR method by combining Dual Cross Pattern (DCP), local binary pattern (LBP) and SVM. Bah SM and Ming F introduced a new FR method [33] by combining LBP and different advanced preprocessing techniques, which is robust under the variation of scale, pose and different lighting conditions.

From different studies, we know that during COVID-19 pandemic, wearing masks helps to prevent the spreading of coronavirus. But masks obscure the important face region and as a result reduce the recognition rate of FR. To increase the FR rate of masked faces in [34] authors presented a Multi-Task Cascaded Convolutional Neural Network (MTCNN) based technique. By combining the convolutional neural network (CNN) and LBP, Vu HN et al. presented a masked FR technique in [35]. In [36], F. Ding et al. presented latent part detection (LPD) model to improve the recognition accuracy of masked faces. Here, the author first generates masked faces and then original and masked faces are fed into two branches CNN. Their technique [36] provides better accuracy compared to others with a large margin. In [37], at first the author cropped the masked face region and then applied CNN, namely VGG-16, AlexNet, and ResNet-50 to extract the features from face regions and then applied Multilayer Perceptron (MLP) for classification.

2.2 Scope and Contribution

The principal contributions of this paper are presented as follows:

1) A CS based novel framework of FR is presented, where CVT performs a dual role:
   (a) Sparses representation using transform domain
   (b) Preprocessing of the face images
2) Proposed method utilizes a new preprocessing technique based on CS, to extract detail edge information from the face images by using CVT which has better ability of providing directional and edge representation.
3) For CS reconstruction Smoothed Projected Land weber (SPL) [38] method has been used for faster implementation.
4) Proposed CS based FR framework improve the recognition rate.
5) Extensive simulations are performed on two data sets AR and ORL. The performance of the proposed method has been evaluated using different classifier such as K-fold cross validation technique, CRC with regularized least square (CRC RLS), NN, Naive Bayes (NB) and SVM classifier.
6) The proposed technique is also a suitable method during the COVID-19 Pandemic.

3 Proposed Method

This section provides the description of proposed CS based FR technique. The structural outline of the proposed FR technique is depicted in Figure 1. It consists of mainly four modules: CVT, CS based preprocessing, PCA for feature selection and classification.

At first CVT [39] has been applied on each input image to capture the detail and directional edge information. Actually CVT is an appropriate basis function for the sparse representation, because maximum numbers of coefficients values are negligible after the application of this transform. So CVT is used in this FR technique due to its high degree of sparsity property compared to other transform. For the reconstruction of image with enhanced information, different percentage of samples (PS) are chosen randomly from the detail sub-bands and then sub-bands are reconstructed back applying SPL [38] reconstruction algorithm. Inverse CVT (ICVT) has been applied on the coarse sub-band and the reconstructed detail sub band. After that we get the resulting image as Image1 (Figure 1).

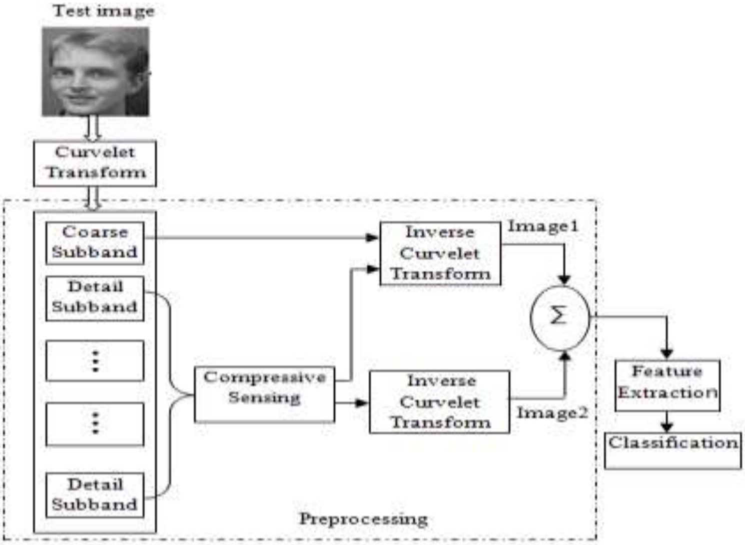


Figure 1
Structural outline of the proposed FR technique

ICVT is also applied on the reconstructed detailed sub-bands only considering all the coefficients of coarse sub-band to zero and then we obtain Image2 (Figure 1). Two images Image1 and Image2 are superimposed to enrich the reconstructed face images in which detail edges are more informative. So the technique enhances the features in this preprocessing.

Both reconstructed images (Image2 and Image1) for different PS are shown in Figure 2. Figure 2 (a) is the input original image, while Figure 2 (b) is the reconstructed Image1 for different percentages of detailed sub-bands, Figure 2 (c) is the reconstructed Image2 from different percentages of detailed sub-bands coefficient only and Figure 2.(d) is the superimposed image. From Figure 2 it is noticed that with the increase in the PS of the detail sub-band coefficients, improvement on reconstruction quality is observed. It is expected that this improvement in reconstruction quality has a subsequent effect on classification rate. PCA has been applied to extract the features from the superimposed images. Then different classifiers are applied on the extracted features to recognize the face images. Algorithm 1 describes the total process of the proposed technique.

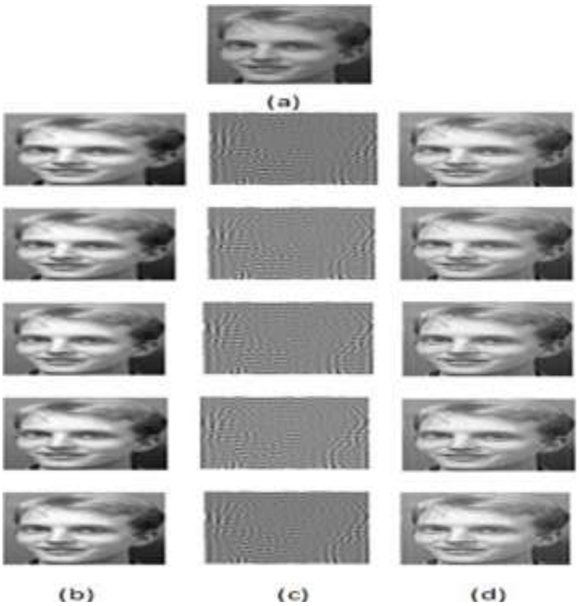


Figure 2

(a)input image, (b)CS based reconstructed images(Image1), by varying the value of PS (c)CS reconstructed images(Image2) for different PS value (d)Superimposed images. (1st row for 50% PS, 2nd row for 60% PS, 3rd row for 70% PS, 4th row for 80% PS, 5th row for 90% PS).

ALGORITHM

Algorithm 1: Algorithm of Face recognition

```
Input: Face image and PS
  1:  Compute CT
  2:  Extract coarse sub-band $Image_{coarse}$ and detail sub-band
      $Image_{detail}$
  3:  Set PS from $Image_{detail}$
  4:  Reconstruct $Image_{detail}$ applying SPL  method
  5:  Construct $Image1=ICT\ (Image_{coarse}, Image_{detail})$
  6:  Set $Image_{coarse} = \mathbf{0}$
  7:  Construct $Image2=ICT\ (Image_{coarse}, Image_{detail})$
  8:  Image = $superimpose(Image1, Image2)$
  9:  Apply PCA
  10: Apply classifier
Output: class label
```

4 Result and Discussions

Every input face image is decomposed using CVT considering scale value of 2 and angle 8. To generate the feature vectors PCA is applied on the enhanced training face images. The feature vectors of the test image are compared with that of the training images to find the best match training image and recognized as the face of the test image. System performance has been evaluated by applying CRC_ RLS [40] and K-fold cross validation technique considering K=10. For K-fold cross validation technique, results are obtained by averaging the recognition rates of 1000 different rounds in MATLAB. Performance of the proposed method is computed on two publicly available facial image databases: ORL and AR. Additionally we have also studied the recognition rate using NN, NB and SVM classifiers. Proposed method is executed on MATLAB 2012b and Weka 3.7.9, in Intel Core i3-380M CPU, 2GB RAM, Windows 7 platform.

4.1 ORL Database

ORL dataset contains grayscale images of 40 individuals with varying illumination, contrast, pose and expressions (open or closed eyes, smiling or no smiling). Some sample images are shown in Figure 3. This database consists of 400 images with frontal and near frontal view faces (rotation of the face up to 20 degrees with and without spectacles).

Recognition rate for different dimensions of feature vector, using cross validation technique is presented in Table 1. Table 2 shows the classification rate for CRC_RLS classifier by varying the PS. It is noticed that recognition rate gradually increases with increase of PS and PC and finally, achieved maximum recognition rate when PC= 50 and PS=90. All the results presented in all tables and graphs are obtained by taking average from 1000 runs.

Figure 3
Few sample images from ORL dataset

Table 1

Recognition rate of ORL dataset for cross validation technique

PS	PC=10	PC=20	PC=30	PC=40	PC=50
90%	87.89%	95.21%	96.43%	96.83%	97.36%
80%	87.75%	95.34%	96.29%	96.74%	97.35%
70%	87.85%	95.33%	96.25%	96.75%	97.28%
60%	87.60%	95.24%	96.28%	96.83%	97.27%
50%	87.77%	95.08%	96.27%	96.69%	97.09%
40%	87.74%	95.14%	96.22%	96.74%	97.19%
30%	87.36%	94.96%	96.35%	96.72%	97.24%
20%	87.38%	94.91%	96.39%	96.73%	97.17%
10%	87.22%	94.88%	96.32%	96.63%	97.25%

Table 2

Recognition rate of ORL database for CRC_RLS

PS	PC=10	PC=20	PC=30	PC=40	PC=50
90%	64%	83 %	87.0%	90.0%	91.5%
80%	64%	82.5%	86.5%	90.0%	91.0%
70%	64%	82%	86.5%	90.0%	91.0%
60%	63.5%	82%	86.0%	90.0%	91.0%
50%	62%	82%	86.0%	89.5%	90.5%
40%	62.5%	81%	85.5%	90.0%	90.5%
30%	62%	80.5%	86.5%	90.0%	90.5%
20%	61%	81%	86.5%	89.0%	90.0%
10%	61%	81%	86.5%	89.0%	90.0%

In this case, in order to observe the dependence of recognition rate on two parameters (feature dimension or PC value and PS value) in the right way, we should trade-off two parameters and the graphical representation will become a three dimensional plot, as shown in Figure 4 and Figure 5. Figure 4 and Figure 5 show how the recognition rate changes according to the changes of the feature dimension and PS value for CRC RLS and cross validation technique. Figure 4 shows that the best recognition rate 91.5% is obtained when the feature dimension is set to 50 and PS= 90%.

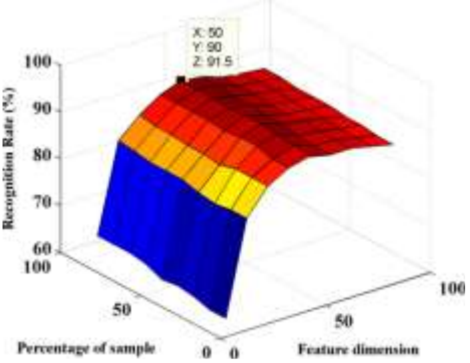


Figure 4
Recognition rate vs. Feature Dimension vs. PS (on ORL dataset considering CRC_RLS)

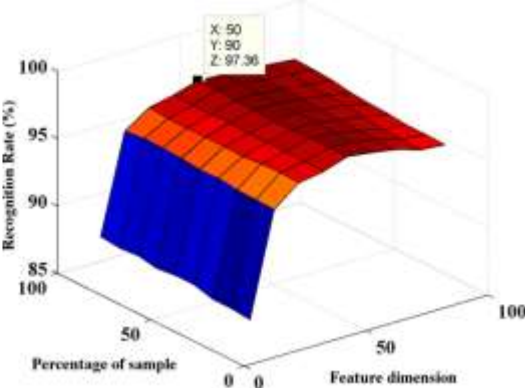


Figure 5
Recognition rate vs. Feature Dimension vs. PS (on ORL database considering cross validation technique)

In case of Figure 5 the maximum recognition rate 97.36% is obtained for feature dimension is equal to 50 and PS=90%. From both the figures it is also observed that the recognition rate decreases even after increasing the PC value after 50. But there is a trend of increase of recognition rate with increase of PS value. Recognition rate for NN, NB and SVM classifiers have been studied by varying the PC and depicted in Figure 6. Maximum recognition rate has been achieved for PC=50, beyond which no significant change in result. For ORL database, SVM and NN classifiers produce the maximum rate of recognition for PC=50. Comparing three classifiers (NN, NB and SVM) from Figure 6, it is seen that SVM provide the best result compared to others for the entire range of feature dimension variation. The ROC curve for NB, NN and SVM classifiers for PS=90% are depicted in Figure 7. Here SVM classifier gives excellent result for the proposed technique.

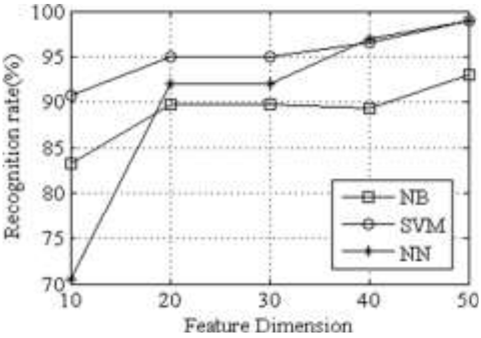


Figure 6
Recognition rate for NB, NN, SVM classifier

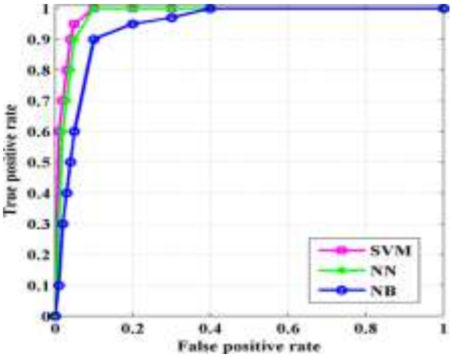


Figure 7
ROC curve for ORL database

The performance of this proposed technique is also compared with the method (CVT+ PCA), where PCA is applied directly on approximation coefficient after applying curvelet transform (CVT is used for curvelet transform). Our proposed method shows superior results than the (CVT+ PCA) as shown in Figure 8. From Figure 8 it has been observed that the proposed method is better than the PCA based methods. For the cases, the best performance has been achieved at PC=50. Results of performance comparison with the existing techniques are summarized in Table 3. The proposed FR technique shows better performance compare to the techniques as described in [27] [42] [41], for PS=90% and PC=50. Computational time required for this proposed pre-processing scheme is given in Table 4.

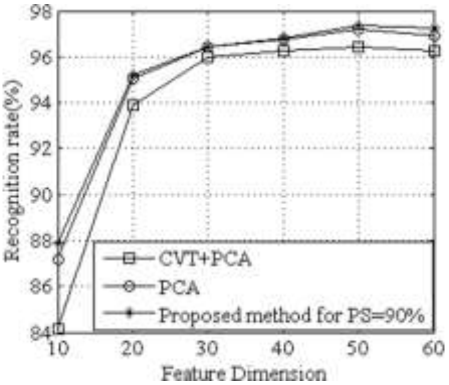


Figure 8
Comparison using cross validation technique

Table 3
Comparisons with other methods for ORL database

Method	Accuracy
PCA+SRC-MP[21]	89.00%(for dim 60)
PCA+SR[41]	93.7%(for dim 100)
Homotopy + SR [42]	97.31%
Proposed Method +SVM	99.00%(for dim 50) 96.00%(for dim 60)
Proposed Method+NN	99.00%(for dim 50) 96.00%(for dim 60)

Table 4
Time required for CS based processing

PS	50%	60%	70%	80%	90%
Time	16.533s	16.633s	16.700s	21.583s	23.894s

4.2 AR Database

The AR dataset consists of 4000 images (consisting of 126 individuals) with variation in illumination and expression. In this work, we choose 1399 images (consisting of 50 males and 50 females) with illumination and expression variation. For each person, 7 images are selected for training and rests of 7 images are used for testing. The images are cropped to (60 × 43) shown in Figure 9.

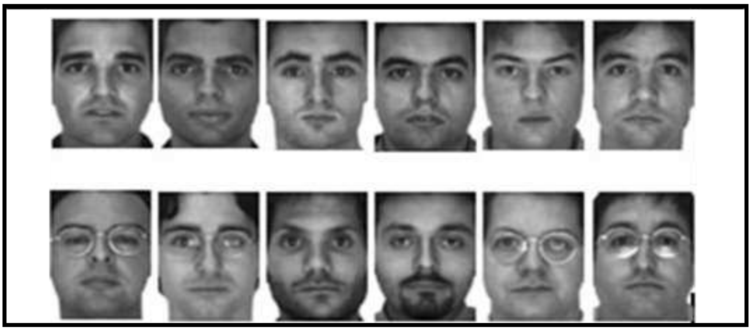


Figure 9
Sample images of AR face database

Recognition rates for different value of PCs and PS using CRC RLS classifier is given in Table 5. For this database also, we have observed the same result that the recognition rate gradually increases with increase of PS and PC. Recognition rates for NN, NB and SVM classifiers have been studied by varying the number of PC and are presented in Table 6 for PS=90%. For AR database, NN classifier produces the maximum accuracy of 98.928% for PS=90%.

Table 5
Recognition rate of AR database using CRC_RLS

PS	PC=60	PC=120	PC=300
90	86.69%	91.99 %	93.99%
80	86.69%	91.41%	93.42%
70	85.69%	91.13%	93.84%
60	85.27%	91.27%	94.28%
50	84.97%	90.55%	93.99%
40	82.97%	90.27%	92.41%
30	83.11%	90.41%	91.56%
20	83.26%	90.41%	91.56%
10	82.69%	90.12%	91.55%

Figure 10 shows how the recognition rate changes according to the changes of the feature dimension and PS value for CRC RLS for AR database. From Figure 10, it is noticed that the best recognition rate 94.91% is obtained when the feature dimension is set to 270 and PS= 90%. From this figure it is also observed that there is a trend of increase of recognition rate with the increase of PS value. The ROC curve for NN and SVM classifiers, using cross validation technique is depicted in Figure 11, showing excellent result. Performance is also compared with PCA (on original image) based methods considering PS=50% and 90% taken from detail sub-band.

Table 6

Recognition rate of AR database for different classifier and PS=90

PC	Accuracy for NB	Accuracy for NN	Accuracy for SVM
60	91.065%	97.856%	97.069%
120	90.278%	98.928%	98.071%
300	83.774%	98.570%	98.000%

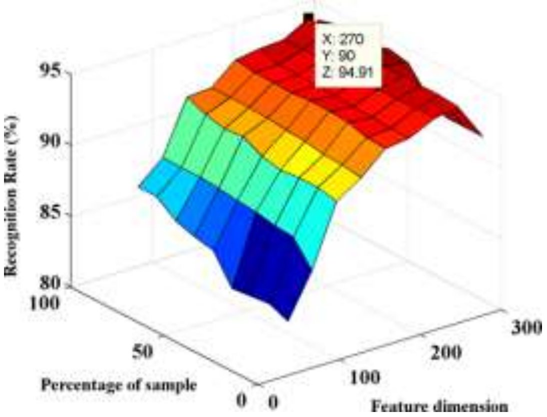


Figure 10

Recognition rate vs. Feature Dimension vs. PS (for AR, considering CRC_RLS)

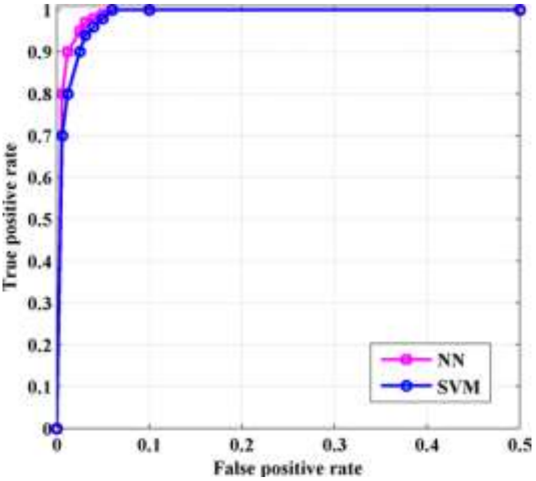


Figure 11

ROC curve for NN and SVM classifier

Comparisons with the existing methods are summarized in Table 7, showing that performance of our technique is better compare to

Table 7

Comparisons with other methods for AR database

Method	Accuracy
PCA+CRC RLS(31)	90.00% (for dim 120)
SRC(21)	90.100% (for dim 120)
Proposed Method+NB	90.278%(for dim 120)
Proposed Method+NN	98.928%(for dim 120)
Proposed method+SVM	98.071%(for dim 120)

5 Modified Proposed Method to Combat COVID19

The modified flow diagram to recognize the face images is depicted in Figure 12.

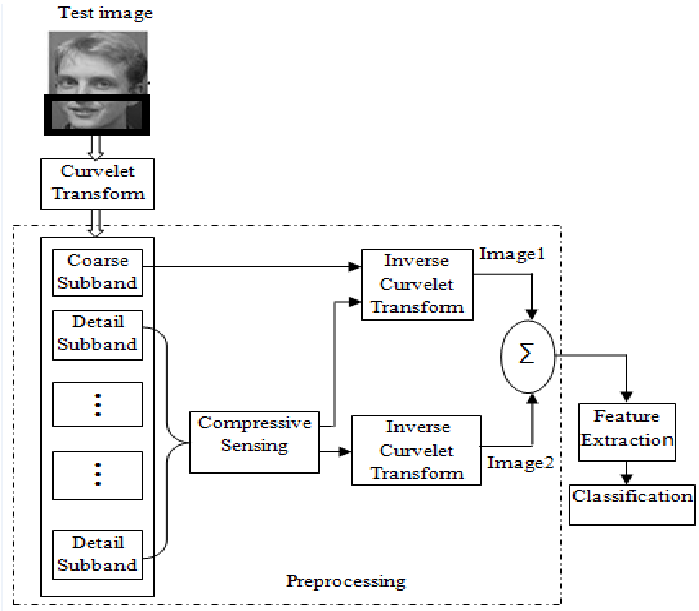


Figure 12

Modified flow diagram of proposed technique to combat with COVID 19

In this technique we cut the 1/3 portion of the face image (as indicated by black box) from the lower region and excluding this lower portion we performed the same technique as discussed in Figure 1. At present, no database is available with a mask and that is why we have used the ORL database. Actually mask covers the lower portion of our face images and as a result we are unable to extract the features of face images, which are covered by mask. So we can extract the features only from the upper uncovered portion. Through this preprocessing technique, as discussed in Figure 1, we have tried to improve the face recognition rate in this pandemic situation to combat with corona virus. Some of sample images of ORL database after excluding the lower portion (which is considered as covered region by mask) are shown in Figure 13. The recognition rate for the ORL database is given in Table 8 for CRC RLS classifier. From the results of Table 8, we can say that the proposed technique performs better than a conventional PCA.

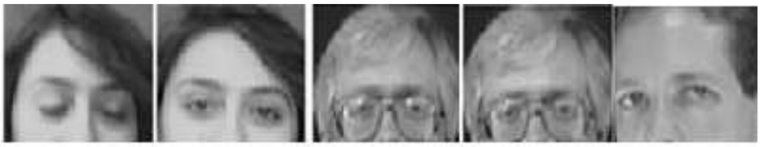


Figure 13

Some of face images of ORL database after excluding the lower portion

Table 8

Recognition rate for ORL database covered with mask

PC	Recognition rate (Conventional PCA)	Recognition rate (Proposed technique)
10	47.00%	51.00%
20	76.50%	79.00%
30	80.00%	82.00%
40	82.00%	86.50%
50	87.50%	89.00%

Conclusions and Future Work

In this paper, a preprocessing technique, based on CS for the performance improvement of FR method, is proposed. This presented integrated FR technique performs preprocessing, compact presentation and dimensionality diminution. The method shows assuring results while the recognition rate is evaluated using CRC RLS, NN, NB and SVM classifiers. Experimental results of this proposed technique show the superiority compared to other methods. Our method shows excellent performances, such as maximum recognition rate of 99% (for SVM classifier) for ORL database and 98.92% (for NN and SVM classifier) for AR database. The proposed CS based FR method, improves rate of recognition and shows robustness against the effect of Gaussian noise. The proposed technique provide maximum of 89% recognition accuracy for the ORL database in case of masked faces, which is greater than conventional PCA. The proposed technique may be extended for future work, as follows:

i) The importance of other feature extraction technique such as LDA and ICA may be studied to improve the rate of face recognition.

ii) A deep learning model can be developed to improve the recognition.

References

[1] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma: Robust Face Recognition via Sparse Representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, No. 2, 2009, pp. 210-227

[2] R. Chellappa, C. L. Wilson, and S. Sirohey: Human and machine recognition of faces: A survey, Proceedings of the IEEE, Vol. 83, No. 5, 1995, pp. 705-741

[3] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, Face recognition: A literature survey, ACM Computing Surveys, Vol. 35, No. 4, 2003, pp. 399-458

[4] A. K. Jain, A. Ross, and S. Prabhakar: An introduction to biometric recognition, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, No. 1, 2004, pp. 4-20

[5] M. Turk, A. P. Pentland, Face Recognition Using Eigen- faces, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1991, pp. 586-591

[6] J. W. Lu, K. N. Plataniotis and A. N. Venetsanopoulos, Face Recognition Using LDA-based Algorithms, IEEE Transactions on Neural Networks, Vol. 14, 2003, pp. 195-200

[7] C. Liu and H. Wechsler: Comparative assessment of independent component analysis (ICA) for face recognition, Proceedings of the Second International Conference on Audio- and Video-based Biometric Person Authentication, 1999

[8] C. C. Liu, D. Q. Dai: Face Recognition Using Dual-Tree Complex Wavelet Features, IEEE Transactions on Image Processing, Vol. 18, No. 11, 2009, pp. 2593-2599

[9] M. N. Do, M. Vetterli: The Contourlet transform: an efficient directional multi resolution image representation, IEEE Transactions on Image Processing, Vol. 14, No. 12, 2005, pp. 2091- 2106

[10] J. L. Starack, E. J. Candes, D. Donoho: Curvelet transform for image denoising, IEEE Transactions on Image Processing, Vol. 11, No. 6, 2002, pp. 670-684

[11] L. Sirovich, M. Kirby: Low-dimensional Procedure for the Characterization of Human Faces, Journal of the Optical Society of America A - Optics, Image Science and Vision, Vol. 4, No. 3, 1987, pp. 519-524

[12] Q. Yang, X. Q. Ding: Symmetrical Principal Component Analysis and its Application in Face Recognition, Chinese Journal of Computers, Vol. 26, 2003, pp. 1146-1151

[13] J. Yang, D. Zhang: Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 1, 2004, pp. 131-137

[14] J. Meng, W. Zhang: Volume measure in 2DPCA- based face recognition, Pattern Recognition Letters, Vol. 28, 2007, pp. 1203-1208

[15] A. P. Kumar, S. Das, and V. Kamakoti: Face recognition using weighted modular principle component analysis, Neural Information Processing, Lecture Notes In Computer Science: Springer Berlin / Heidelberg, Vol. 3316, 2004, pp. 362-367

[16] V. D. M. Nhat, S. Lee: An Improvement on PCA Algorithm for Face Recognition, Advances in Neural Networks, Lecture Notes in Computer Science. Chongqing: Springer, Vol. 3498, 2005, pp. 1016-1021

[17] D. Zhang, Z.-H. Zhoua, S. Chen: Diagonal principal component analysis for face recognition, Pattern Recognition, Vol. 39, 2006 pp. 140-142

[18] Deeba Farah, Memon Hira, Ali Dharejo Fayaz, Ahmed Aftab, Ghafar Abddul: LBPH based enhanced real-time face recognition, Int J Adv Comput Sci Appl.10(5), 2019

[19] Nisha Maitreyee Dutta, Improving the recognition of faces using LBP and SVM optimized by PSO technique, Int J Exp Diabetes. 5(4), 2017, pp. 297-303

[20] M. Yang, l. Zhang: Gabor Feature based Sparse Representation for Face Recognition with Gabor Occlusion Dictionary, Proceedings of the 11th European conference on Computer vision, 2010, pp. 448-461

[21] M. Yang, L. Zhang, J. Yang, D. Zhang: Robust Sparse Coding for Face Recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 625-632

[22] R. He , W. S. Zheng, and B.G. Hu: Maximum Correntropy Criterion for Robust Face Recognition, IEEE Transactions on Biometrics Compendium, Vol. 33, No. 8, 2011, pp. 1561-1576

[23] J. Huang, X. Huang, and D. Metaxas: Simultaneous image transformation and sparse representation recovery, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1-8

[24] Z. Zhou, A. Wagner, H. Mobahi, J. Wright, and Y. Ma: Face Recognition With Contiguous Occlusion Using Markov Random Fields, Proceedings of the IEEE International Conference on Computer Vision, 2009, pp. 1050-1057

[25] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, Y. Ma: Towards a practical face recognition system: robust registration and illumination by sparse

representation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 597-604

[26] M. Yang, L. Zhang, X. Feng, and D. Zhang: Fisher discrimination dictionary learning for sparse representation, Proceedings of the IEEE International Conference on Computer Vision, 2011, pp. 543-550

[27] Yi-Haur Shiau1 and Chaur-Chin: A Sparse Representation Method with Maximum Probability of Partial Ranking For Face Recognition, Proceedings of the IEEE International Conference on Image Processing, 2012, pp. 1445-1448

[28] J. Wang, C. Lu, M. Wang, X. Hu: Robust Face Recognition via Adaptive Sparse Representation, IEEE Transactions on Cybernetics, Vol. 44, No. 12, 2014

[29] Cheheb I, Al-Maadeed N, Al-Madeed S, Bouridane A, Jiang R: Random sampling for patch-based face recognition, 5th international workshop on biometrics and forensics (IWBF), IEEE, 2017, pp. 1-5

[30] Iliadis M, Wang H, Molina R, Katsaggelos AK: Robust and low-rank representation for fast face identification with occlusions, IEEE Trans Image Process 26(5), 2017, pp. 2203-2218

[31] Zhao F, Feng J, Zhao J, Yang W, Yan S: Robust LSTMautoencoders for face de-occlusion in the wild. IEEE Trans Image Process 27(2), 2017, pp. 778-790

[32] Bhangale Kishor B, Jadhav Kamal M, Shirke Yogesh R: Robust pose invariant face recognition using DCP and LB, International Journal of Management, Technology and Engineering September 2018;8(IX):1026–34. ISSN NO: 2249-7455

[33] Bah SM, Ming F: An improved face recognition algorithm and its application in attendance management system. 2020, Array 5:100014

[34] M. S. Ejaz and M. R. Islam: Masked Face Recognition Using Convolutional Neural Network, 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), 2019, pp. 1-6

[35] Vu HN, Nguyen MH, Pham C: Masked face recognition with convolutional neural networks and local binary patterns. Appl Intell (Dordr). 2021 Aug 14, pp. 1-16

[36] Feifei Ding, Peixi Peng, Yangru Huang, Mengyue Geng, and Yonghong Tian: Masked Face Recognition with Latent Part Detection, Proceedings of the 28th ACM International Conference on Multimedia Association for Computing Machinery, New York, NY, USA,2020, pp. 2281-2289

[37] Walid Hariri: Efficient masked face recognition method during the COVID-19 pandemic, Signal, Image and Video Processing, November, 2021

[38] S. Mun and J. E. Fowler: Block Compressed Sensing of images using Directional Transforms, Proceedings of the international conference on Image processing, 2009, pp. 3021-3024

[39] E. J. Candes, L. Demanet, D. L. Donoho, L. Ying: Fast discrete curvelet transform, SIAM Journal of Multiscale Modeling and Simulations, Vol. 5, No. 3, 2007, pp. 861-899

[40] L. Zhang, M. Yang, X. Feng: Sparse representation or collaborative representation: Which helps face recognition?, Proceedings of the IEEE International Conference on Computer Vision, 2011, pp. 471-478

[41] Y. Wang, C. Wang and L. Liang: Sparse Representation Theory and Its Application for Face Recognition, International Journal on Smart Sensing and Intelligent Systems, Vol. 8, No. 1, 2015

[42] Z. Zhang, Y. Xu, J. Yang, X. Li and D. Zhang: A Survey of Sparse Representation: Algorithms and Applications, IEEE Access, Vol. 3, 2015, pp. 490-530

Business Culture and Behavioral Characteristics

Zsuzsanna Tóth, László Józsa, Erika Seres Huszárik

J. Selye University, Bratislavská cesta 3322, 945 01 Komárno, Slovakia
e-mails: tothz@ujs.sk; jozsal@ujs.sk; huszarike@ujs.sk

Kim-Shyan Fam

Széchenyi István University, Egyetem tér 1, H-9026 Győr, Hungary
e-mail: kimfam@magscholar.com

Abstract: The main goal of our research, and thus, of our present study, was to explore some problems and issues of business behavior and etiquette in Slovakia and Hungary. The international comparative research program launched by Fam and Richards was our starting point, in which we examined these two countries. We found that due to the cultural differences in the dimensions of the Hofstede model, differences can be detected in business ethics and etiquette in the business life of Hungary and Slovakia, which can be supported by statistical methods. At the same time, our results also showed that almost a half-century since Hofstede research has not passed without a trace in the Central European Region. The transition from socialism to a market economy involved border openings. At the same time, it facilitated the convergence of the business culture of Slovakia and Hungary, changing the relative position of these two countries on the Hofstede scale. We drew attention to the fact that it would be worth repeating Hofstede's research to record socio-economic changes, in the case of intensely transforming societies and countries.

Keywords: business behavior; business etiquette; business ethics; cultural differences

1 Introduction

Culture largely determines people's daily lives, about which many definitions have come to light over the past decades. Culture is the same age as humanity. The complexity of the concept reflects the fact that while Alfred Kroeber and Clyde Kluckhohn compiled 164 definitions of culture in 1952, by now, this number has presumably reached the order of a thousand [39]. Although all these definitions are close to each other, they differ. They depend on the age and society, the approach, and the purpose of viewing culture [46]. The purpose of the present study is to map, following Hofstede's culture model, how to behave in a business meeting, the patterns of behavior accepted by business people, and the key to a successful business in Slovakia and Hungary. On a theoretical level, therefore, we review the concept and role of culture in business and then use the results of primary research to present the similarities and differences between the business cultures of the two countries.

1.1 The Concept of Culture

Culture is the unique nature of a social group that distinguishes it from other social groups. Culture develops from patterns of behavior created by a group of people in response to fundamental problems of social interaction. It manifests in the values, beliefs, and norms of a group, the typical patterns of behavior of group members, the choice and use of rituals and symbols, the social, economic, political, and religious institutions, and the ideology that underlies the institutions.

Perhaps the most significant cultural research is Geert Hofstede's research, but Trompenaars's study [57], examining cultural values is significant. Since the publication of Culture's Consequences: International Differences in Work-Related Values [27], thousands of empirical studies have been inspired by this work [51]. According to the Social Science Citation Index, Hofstede's work is more widely cited than other studies. The most important cultural models are: Hofstede [27], Hofstede and Bond [29], Hofstede [28], Schwartz [54], Trompenaars [57], Smith et al. [56], House et al. [32], Bond et al. [7], McLean and Lewis [44].

Hofstede [27] interprets culture as the collective programming of thinking that distinguishes one group of people from others. He developed his theory based on his research - data collected from 116,000 questionnaires in 20 languages, involving 88,000 employees in 72 countries at IBM between 1967 and 1969, and again between 1971 and 1973.

The *individualism/collectivism* factor expresses the extent to which individuals care only for themselves and their close family and how much they feel responsible for members of a large community who can also count on their support in return. The factor of *avoiding uncertainty* expresses how members of a community can face uncertainty and take risks. Three indicators play a role here: adherence to the rules, duration of employment, and stress endurance. *Power distance* is an expression of the extent to which members of a society or community who receive less power accept the unequal distribution of power. *Masculine / feminine* values refer to gender-related role sharing in a given society. For example, East Asia, Central Europe, and the Anglo-Saxon states are predominantly masculine societies. In contrast, northern and Latin Europe, and many African cultures, show more notable feminine characteristics.

Any of Hofstede's four dimensions can have an impact on the methods used in negotiations. In addition, each dimension can affect the relationship between the negotiators. As a result, all of these can shape the negotiation process and its outcome [23].

Hofstede and Bond [29] created the fifth dimension, Confucian dynamism (*long-term/short-term orientation*). Long-term orientation refers to future-oriented values such as perseverance and frugality. Short-term orientation refers to past and present values such as respect for traditions and fulfilment of social obligations.

Later, the sixth dimension was added to Hofstede's [27] model, *Indulgence versus Restraint*. The indicator refers to the level of acceptance of each culture related to the enjoyment of life and entertainment, or how restrained each culture is due to strict social norms [28].

Hofstede's [27] work has been widely criticized, among other things, for grouping culture into four to five dimensions, restricting sampling to a multinational firm, or neglecting cultural heterogeneity within a country [53]. However, despite criticisms, researchers prefer this kind of 5-dimensional division because of its clarity and providence [35].

Hofstede's dimensional concept of culture dominates in international management and cross-cultural psychology; on the other hand, his dimensional concept neglects cultural dynamics [6]. With the economic development of countries, modernization theory predicts changes in cultural values. In Hofstede's value dimensions, country scores may change, raising the further relevance of the framework [5]. Eringa et al. [17] and Gerlach and Eriksson [21] repeated some elements of Hofstede's research. They observed significant differences from the original model in several respects, suggesting that individual cultures are not constant but constantly changing.

1.2 The Role of Ethics in Business

In today's globalized world, business actors face several ethical questions in their daily decisions in a dynamic and changing environment. The role of ethics in business has opened up a remarkably new field of research. Most bibliometric studies focus on the volume and citation of papers on the subject [8] [9]. In business situations, the question often arises in the minds of actors: What should I do? What is the right thing to do? In answering this question, the individual's business principles and personal values and emotional intelligence, which influence daily life and social relationships [41], all play a role [19].

The papers published on this topic in the last decades can be divided into four groups. The first group examines the degree of the ethicality of the individual in the context of entrepreneurial skills: e.g. the role of personal values [26], socio-cultural background [31]. Hannafey [25] points to the unique and diverse moral problems and ethical dilemmas entrepreneurs often face. Other research examines organizational structure and the evolution of ethics [55], focusing on the relationship between corporate strategy and values that favor (un)ethical behavior [47]. It is typical of the studies in this group that the authors point to differences in ethical behavior between entrepreneurs and non-entrepreneurs, emphasizing the differences in behavior between managers and entrepreneurs [13] [38]. Crane [13] investigated the behavioral differences between managers and entrepreneurs in Canada and found slight differences in ethical behavior between the two groups. Zhang and Arvey [62] published a fascinating study on whether there is a link between the ethical business behavior of entrepreneurs who are highly rule-breaking in adolescence and their ethical business behavior later in adult life.

Another vital contribution to this area is the Sackey, Faltholm and Ylinenpaa [52] study, which pointed out the ethical dilemmas in developed and developing countries. The authors pointed out that the ethical difficulties faced by entrepreneurs in developed countries are substantially different from those faced by entrepreneurs in developing countries, as business actors in the two countries face other challenges. The research experience shows that in linking the concepts of ethics and entrepreneurship, it is essential to emphasize the moral constraints of entrepreneurs. There is also extensive literature on this area of research [12] [63].

There is ample evidence in the literature that the personal characteristics of entrepreneurs make them sensitive in preparing ethical decisions. In this respect, Pellegrini and Ciappei [49] suggest that an individual's skill enables them to make the right decision even in extreme situations with high uncertainty. Another strand of research on ethical decision-making focuses on non-ethical decision-making. Researchers try to find answers to the personal motives behind individuals' unethical decisions. Baron, Zhao, and Miao [4] found a link between money-driven motivation and moral apathy and concluded that moral indifference predisposes to unethical choices.

In the second group, we can classify papers that examine the issue of ethics at the organizational level, i.e. how ethics is manifested in established organizations and how ethics can evolve within the organization as the company's life cycle progresses. Arend's [3] study points out that the organization's dynamic capabilities actively change existing ethical concepts. The results also indicate a positive impact on the overall ethical performance of the organization. Researchers in the field examine their critical findings in the context of stakeholders [45] [20], and corporate social responsibility [24] [33] [40] [58]. Markman, Russo, Lumpkin, Jennings, and Mair [43] point out the importance of one of the roles of organizations to positively impact their environment and society as a whole. The authors show that there are many ways for organizations to pursue a sustainable, ethical and entrepreneurial strategy simultaneously. It is increasingly apparent that organizations today are becoming more and more committed to corporate social responsibility [18].

The third group includes papers that discuss ethical issues of new business models, e.g. social enterprises [61] [22]. Literature studies agree that social enterprises create social value. At the heart of the operation of these enterprises is a response to a social issue [37]. Despite the growing attention that researchers are paying to social entrepreneurship, few have explored the ethical context in which it operates. Instead, research findings only highlight the social and general economic differences between these types of enterprises [10]. Kraus et al. [36] and Rey-Martí et al. [50] use a bibliometric approach, to provide a comprehensive picture of the outcomes of social purpose enterprises, making clear the importance of the topic. Understanding the ethical principles behind the operation of social enterprises raises further questions in the minds of researchers [10]. Chell et al. [10] point out that social entrepreneurship needs to be viewed through the 'mirror' of ethics and that there is currently no successful integration of the two concepts. To fill this gap, they emphasize the importance of considering ethical perspectives in social entrepreneurship. The authors argue that a positive understanding of social enterprises is superficial simply because they contribute to the common good since there are also fundamental business interests behind these enterprises. Dey and Steyaert [15] are also critical of the ethics of social entrepreneurship. They call for further research to be conducted in this area, with a particular focus on rethinking ethical approaches used in the past. They argue that future research might be worthwhile within the field to investigate issues of power-seeking, subjectivity, and the individual's desire for freedom.

Finally, the last group of literature includes papers that examine the broader perspective of ethical business and its impact on society. For example, the research of Anokhin and Schulze [2], using sources from 64 countries, examines the effects of corruption in the corporate environment, summarizing its impact on society. In line with each other, Kaback [34], Pearson, Naughton and Torode [48] and Von Schnitzler [59] have looked at the ethical issues involved in the introduction of new technology into society. Collewaert and Fassin [11] investigated the effects of unethical behavior on the origins and course of conflicts. In their study, they concluded that what business partners perceive as unethical behavior leads to conflict. The authors also conclude that the disputes described above influence the choice of future business partners and the development of business strategies. A further study on the subject has been carried out on investors [16]. The authors investigated how an investor's reputation affects the success of an investment. Their results show that an ethically questionable decision made in the past can even lead to the rejection of a potential business relationship. Their research suggests that the investor's poor ethical reputation can significantly undermine the added value provided by a partnership and the past success of an investor. The issue of entrepreneurial ethics is becoming increasingly important, particularly in emerging and developing economies [1], [14] [60]. Cumming et al. [14] point out that businesses that follow ethical behavior contribute significantly to the development of the Chinese economy and poverty reduction in Chinese society.

As a result of our literature review, we discovered that there are still several unanswered questions in this research area. One of these is the different behavior of individuals and decision-makers in business situations, which we have studied, resulting in a series of ethical and unethical decisions and actions.

2 The Aim of the Study and the Applied Research Method

We aimed to synthesize and summaries the scientific results that deal with the specifics of business behavior, especially concerning individual cultures, and examine companies operating in Slovakia and Hungary in terms of business ethics and etiquette through theoretical foundations. To achieve our goal, we have identified our main research question:

> **Are there any differences or peculiarities in the field of Slovak and Hungarian business ethics and etiquette that can be discovered?**

Our hypothesis is as follows:

> **H1:** Due to the cultural differences in the dimensions of the Hofstede model, despite the common historical past, we can discover significant differences in the business life of Hungary and Slovakia in the field of business ethics and etiquette.

The empirical research was based on an online questionnaire survey, and its participants were representatives of Hungarian and Slovak companies. The questionnaire examines the participants' behavior. This research method also has disadvantages, i.e., the respondents may not be willing and able to provide accurate information to the questions asked. In addition, answering personal and sensitive questions can be a disadvantage [42].

The questionnaire query was followed by data cleansing and evaluation. To examine our hypotheses, we chose the primary research method, including the one-time, descriptive analysis, especially as we obtained our data on one sample at a time [42].

Our empirical research was carried out as part of a more extensive international research project. Within the framework of an international project - Marketing in Asia Group, New Zealand - Slovakia and Hungary were examined in terms of business communication, ethics and etiquette. Professor Kim-Shyan Fam and Dr. James E. Richard, research leaders from Victoria University of Wellington, compiled and tested the questionnaire. The available questionnaire was translated from English into Hungarian and Slovak. Then an independent translator, with no prior knowledge of the original content, translated it back into the original language to allow an accurate cross-cultural comparison.

To collect the data, we needed to create a database of companies operating in Slovakia and Hungary. The size of the companies and the industry were not dominant. The compiled mailing list using the collection pages included the contact details of 938 companies. We sent out our online questionnaire in the spring of 2018. Due to invalid, non-functioning e-mail addresses, our e-mail failed to be delivered in 22 cases. In the course of the research, we used the snowball method from the random sampling procedures. We collected our company manager acquaintances, to whom we forwarded our online questionnaire and asked them to pass it on to managers. After data cleansing, we had a total of 257 completed questionnaires, so the willingness to respond in three months was 28.05%. From this, we can conclude that respondents are likely to be reluctant to participate in such surveys. A total of 103 respondents from Hungary and 154 from Slovakia participated in our research. The completed questionnaires were coded, and then the obtained values were recorded in the table of the SPSS statistical program. This program also conducted the evaluation: univariate, bivariate, and multivariate analyses were performed.

It was necessary to review the cultural dimensions set up by the Dutch social psychologist Geert Hofstede to examine the hypotheses thoroughly and carefully. Hofstede classifies national cultures along six dimensions to characterize society's nature carefully. The most significant difference between Slovakia and Hungary is in the power distance indicator. Slovakia is characterized by an exceptionally high level of power distance between its leaders and their subordinates. Slovakia's masculinity indicator, which measures the cultural prevalence of strength and competition, is also higher than in Hungary. Still, the difference, in this case, is smaller than in the power distance indicator. Interestingly, Slovakia is at the forefront of the world, in these two indicators.

According to Hofstede's cultural classification, Slovak society is also more future-oriented than Hungarian society. Hungary is ahead of Slovakia in the indicators of individualism and avoidance of uncertainty. The most significant difference is between the power distance indicator of Hungary and Slovakia, as Slovakia here has a value of 104 and Hungary has 46. Power distances significantly affect business, including business ethics and business etiquette. In terms of individualism, Hungary scored 80 and Slovakia had 52. The location of culture on the individualist-collectivist axis also has a severe impact on business relations. In the indicator of masculinity related to competition and aggression, Hungary has an exceptionally high score of 88, but Slovakia has an outstanding value of 110. The competitive situation also affects the course of business relations. It is essential to take the right amount of risk in business and plan for the future correctly. In both cases, Slovakia excels with its lower uncertainty avoidance score and higher future orientation value. These are certainly reflected in business relationships. Unlike the other dimensions, leniency does not show a significant difference between the two countries.

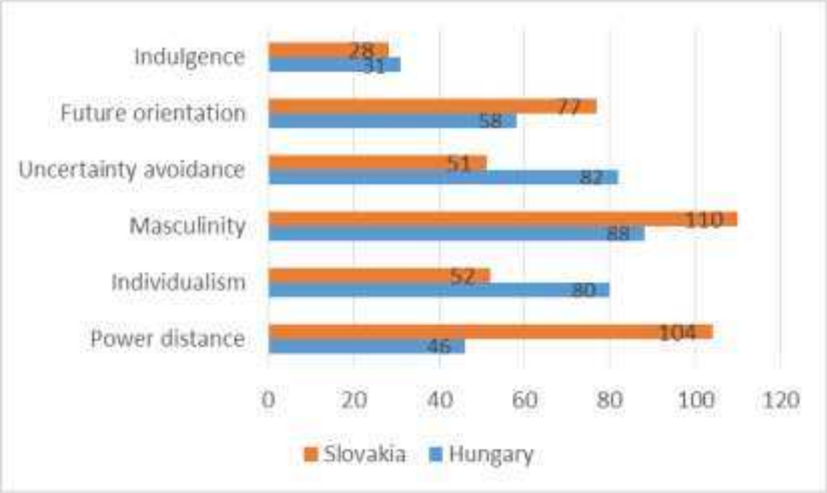


Figure 1

Values of Hungary and Slovakia in the cultural dimensions of Hofstede

Source: Values of Hungary and Slovakia in the cultural dimensions of Hofstede [30]

Respondents on a seven-point scale rated each element of business etiquette. It is possible to explore possible differences between the responses in Hungary and Slovakia by comparing the averages within the group. The relevant statistical test, the t-test, shows whether the two means are the same, i.e. whether we should reject the null hypothesis that their difference is not statistically different from zero. After performing the t-test, it can be stated that there is no statistically significant difference between Slovakia and Hungary in terms of personal appearance, professional behavior and social behavior. At the 10% significance level, it can be shown, that compared to the Slovak respondents, communication, cultural sensitivity, accuracy, respect, trust, and reciprocity are somewhat more important for the Hungarian respondents. Our respondents in Slovakia consider only gift-giving significantly more vital than those in Hungary.

Table 1

Differences between Slovakia and Hungary in some elements of business etiquette (N = 241)

	Slovakia	**Hungary**	**Difference**	**p-value of t-statistic**
Communication	5.38	6.07	-0.69	0.000
Cultural sensitivity	4.74	5.04	-0.30	0.066
Gift-giving	4.94	4.40	0.53	0.002
Personal appearance	5.92	5.77	0.15	0.302
Professional behavior	5.92	6.00	-0.08	0.580
Punctuality	5.84	6.11	-0.27	0.070
Respect	5.75	6.26	-0.50	0.000

Social behavior	5.66	5.83	-0.17	0.239
Trust	5.59	6.23	-0.64	0.000
Reciprocity	5.26	5.85	-0.59	0.001

Source: Author's editing

The obtained result somewhat contradicts the difference revealed in the comparison of cultural dimensions between the two countries. Respondents in Slovakia feel that gift-giving is significantly more important than respondents in Hungary do. Presumably, gift-giving is essential in Slovakia because it makes it possible to bridge significant distances of power, but it is also conceivable that the collectivist nature of Slovak society is causing this phenomenon. This element of business etiquette is, therefore, more present in the Slovak business culture.

Many business ethics issues can be related to a country's business culture, which can be classified using Hofstede's cultural dimensions. The business people participating in the questionnaire had to evaluate the conditions in Slovakia and Hungary on a seven-point scale according to the extent to which elements of business ethics prevailed in their most recent business transaction, which is in the early stages of the business relationship. The averages in Slovakia and Hungary moved roughly together in the sample. Determining whether there is a statistically significant difference between the two countries can be done with a t-test. The results of the t-test confirm the similarities predicted by the means: in most cases, there is no significant difference between Slovakia and Hungary; the results cannot be considered statistically different.

Table 2

Differences between Slovakia and Hungary in some elements of business ethics in the initial stage of business relations (N = 52)

	Slovakia	Hungary	Difference	p-value of t-statistic
Business transparency	5.24	5.56	-0.32	0.270
Commitment to the business relationship	4.88	5.67	-0.78	0.031
Credibility	5.09	5.89	-0.80	0.054
Equal opportunities	4.59	4.89	-0.30	0.413
Fair competition	4.76	5.17	-0.40	0.245
Justice (general)	4.91	5.22	-0.31	0.441
Management transparency	4.74	5.11	-0.38	0.337
Sincerity	5.09	5.39	-0.30	0.458
Integrity	5.18	6.00	-0.82	0.066
Keeping promises	5.29	5.50	-0.21	0.540
Loyalty to the relationship	4.97	5.89	-0.92	0.039
Reliability	4.76	5.56	-0.79	0.061

Respect	4.76	5.28	-0.51	0.236
Similar morals	4.91	5.28	-0.37	0.460
Social responsibility	4.68	5.00	-0.32	0.426

Source: Author's editing

At the same time, there was a significant difference in favour of Hungary regarding commitment and loyalty to the business relationship, reliability, integrity and credibility. The higher value of reliability and integrity can also be explained by the fact that Hungary is more avoidable of uncertainty than Slovakia.

Business etiquette consists of many elements. Assessing these, especially concerning the initial phase of the business relationship, was also the task of the survey participants from both countries. In the first group of relevant questions, the result was remarkable: all the listed characteristics (bilingualism, formal communication, direct speech, cultural tolerance, respect for hierarchy) were considered more important in Hungary than in Slovakia. The significance of the differences between one and one and a half units on the seven-point scale is also statistically supported by the t-test. The importance of respecting the hierarchy would have been more expected in Slovakia, a high power distance country, so this result is contrary to expectations. At the same time, compared to Slovakia, formal communication, cultural tolerance, and direct speech are more important, which may reflect the less masculine Hungarian society. The value of bilingualism can be outstanding because the average language skills in Hungary are poor, so multilingualism is rather prominent. There are far fewer differences for the next set of questions assessing business etiquette in the initial phase than the previous data.

Table 3

Differences between Slovakia and Hungary in some aspects of business etiquette in the initial stage of business relations (N = 52)

	Slovakia	**Hungary**	**Difference**	**p-value of t-statistic**
Recognition of hierarchy	5.00	5.33	-0.33	0.172
Assessing cultural differences	5.50	5.50	0.00	1.000
Cultural adaptability	5.26	5.83	-0.57	0.079
Preserving prestige, attending meetings. accepting invitations	5.35	5.44	-0.09	0.721
Gift-giving is required	4.74	4.61	0.12	0.758
Awareness of social status	4.65	4.61	0.04	0.933
Use of titles, qualifications	4.32	4.22	0.10	0.841
Providing appropriate expensive gifts	4.41	3.56	0.86	0.110
Appropriate dressing	5.24	5.00	0.24	0.492
Live-to-work attitude	4.91	4.78	0.13	0.739

Mutual trust is the key to success	5.53	5.72	-0.19	0.503
Commitment to the relationship	5.68	5.61	0.07	0.838
Mutual trust	5.85	6.06	-0.20	0.546
To be competent	4.47	5.50	-1.03	0.046
To show loyalty	4.53	5.56	-1.03	0.042

Source: Author's editing

After performing the t-test, the only statistically significant differences appear in loyalty, competence and cultural adaptability. These were considered more important by the Hungarian respondents. The difference in the perception of competence and loyalty cannot be explained clearly by the indicator of power distance. It is higher in Slovakia, but somewhat yes with the higher individualism in Hungary. Cultural adaptability is challenging to reconcile with the strength and competitive nature of higher masculinity in Slovakia, which may have contributed to the outcome.

The following relevant set of questions in the survey also reveals several similarities in the business etiquette of the two neighboring countries.

Table 4
Differences between Slovakia and Hungary in some aspects of business etiquette in the initial stage of business relations (N = 52)

	Slovakia	**Hungary**	**Difference**	**p-value of t-statistic**
Addressing people with their proper titles	5.35	5.44	-0.09	0.808
Exaggeration	4.74	3.61	1.12	0.021
Confidentiality	5.26	5.28	-0.01	0.976
Fulfilment of obligations	5.59	6.11	-0.52	0.157
Providing appropriate solutions	5.65	5.78	-0.13	0.729
Relationship and business transparency	4.94	5.33	-0.39	0.320
Punctuality	5.21	6.00	-0.79	0.097
Strong handshake	5.32	5.61	-0.29	0.508
Maintaining harmony	5.18	5.67	-0.49	0.203
Respect for all parties	5.18	6.00	-0.82	0.017
Assessing and preserving authority	5.29	5.61	-0.32	0.331
Different attitudes towards authorities	5.44	5.61	-0.17	0.605
Developing personal relationships	5.21	6.11	-0.91	0.027
Great host	5.82	5.89	-0.07	0.828
Preserving humor	6.06	5.78	0.28	0.344

Source: Author's editing

After completing the t-test, statistical analysis, exaggeration, accuracy, development of personal relationships, and respect for all parties are the elements in which Slovakia and Hungary differ. Based on the responses, respect for all parties and developing personal relationships in Hungary seem more critical. This result can be explained by the less masculine, less competitive cultural environment in Hungary. In Slovakia, on the other hand, the role of exaggeration is more significant. The reason is that society reflects the masculine features of power more in Slovakia than in Hungary.

Many dimensions of business etiquette also include how acceptable, inappropriate, or even appropriate certain behaviors are in a country's business. The respondents also had to answer relevant questions, keeping in mind the initial stage of the business relationship. Respondents rated on a seven-point scale how appropriate or even incorrect the six types of behavior listed were:

- Face-to-face encounters
- Direct speech
- Gossip about the customer
- Use of aggressive sales tactics
- Using only the first name in the introduction
- Direct communication

Interestingly, the respondents in Hungary and Slovakia took a similar position on specific issues, while there were even striking differences in other cases. Statistical, formal testing of any discrepancies between the two countries can be performed using a t-test.

Table 5

Differences between Slovakia and Hungary in the assessment of the appropriateness of certain types of behavior in the initial stage of business relations (N = 52)

	Slovakia	**Hungary**	**Difference**	**p-value of t-statistic**
Face-to-face encounters	4.71	5.89	-1.18	0.020
Direct speech	4.82	6.17	-1.34	0.013
Gossip about the customer	3.06	2.83	0.23	0.651
Use of aggressive sales tactics	4.00	2.89	1.11	0.055
Using only the first name in the introduction	3.50	3.33	0.17	0.733
Direct communication	5.41	5.89	-0.48	0.135

Source: Author's editing

The results proven by the t-test show that there is no statistically significant difference between the two countries in assessing the appropriateness of first-name introduction, direct communication, and gossiping. For business people in

Hungary, the form of behavior characterized by face-to-face encounters and direct speech seems to be significantly more appropriate than in the case of Slovak respondents. In contrast, aggressive sales tactics were statistically significantly more appropriate in Slovakia, than in Hungary. All this is in perfect agreement with the fact, established on the basis of Hofstede's cultural dimensions, that masculine traits such as force are more accepted in Slovak society. Consequently, its use in sales may not seem unacceptable either.

The questionnaire survey results among the Hungarian and Slovak respondents outlined above show that not in all respects, but in many cases, differences in business ethics and etiquette supported by statistical methods can be detected between Hungary and Slovakia. Some of the differences revealed seem to contradict some of the cultural differences based on the dimensions of the Hofstede model, but most of them reflect cultural dimension values. Thus, the differences demonstrated in the answers to the questionnaire can, in many cases, be explained, among other things, by uncertainty avoidance or the collectivist-individualist distinction. Among the explanatory cultural dimensions, masculinity stands out, in which, according to Hofstede's classification, Slovakia has a higher score than Hungary. Strength, the prevalence of competition and their social acceptance are thus more significant in Slovakia than in Hungary, which may be the reason for several identified differences. For example, aggressive sales techniques should be emphasized because, from a statistically significant point of view, the Slovak respondents consider it more acceptable than respondents in Hungary.

Based on the performed analysis, we can state that we can accept hypothesis H1 of our research, according to which cultural differences in the dimensions of the Hofstede model can in many cases reveal differences in business ethics and etiquette in the business life of Hungary and Slovakia.

Conclusion and Managerial Implications

Hofstede's cultural model has already drawn attention to the fact that, despite their geographical proximity, significant differences can be detected between the cultural dimensions of the two countries included in our study. Slovakia and Hungary differ primarily, in terms of power distance and uncertainty avoidance. While Slovakia is more characterized by significant power distance, Hungary has a higher value in avoiding uncertainty. While our respondents in Slovakia mostly use gift-giving to bridge the power distance, our Hungarian companies value their commitment and loyalty to business relationships more because of uncertainty avoidance. On the other hand, based on our research, someone in Hungary is more favorable if they are accurate, respects their partner, is committed, loyal and reliable. Based on the above, we can state that if a new economic player wants to enter the market of the two countries, it is worth preparing for the first meeting in Slovakia with a smart gift-giving business strategy and even using aggressive sales techniques.

Our research confirmed the Hofstede model and our initial expectations that there are significant differences in the business culture and etiquette of Slovakia and Hungary and that these differences can be unambiguously demonstrated by appropriate research methodology and statistical analyzes. Conversly, the differences can be explained by the era of socialism, where the role of the planned economy and corporate independence was different, and on the other hand by the different ways of regime change, according to which different managerial cultures gained ground in the two countries. Our results also showed that the near half-century since Hofstede's research has not passed unnoticed in Central Europe. The transition from socialism to a market economy brought with it the opening of borders. At the same time, it facilitated the convergence of the business culture of Slovakia and Hungary, changing the relative position of these two countries on the Hofstede scale. Although we cannot clearly state it due to the limitations of our research, we would like to draw attention to the fact that it would be worth repeating Hofstede's research, to record socio-economic changes in the case of dynamically changing societies and countries.

References

[1] AHMAN, N. - RAMAYAH, T. Does the notion of ‘doing well by doing good’ prevail among entrepreneurial ventures in a developing nation? *Journal of Business Ethics*, 106 (4): 479-490, 2012

[2] ANOKHIN, S. - SCHULZE, W. S. Entrepreneurship, innovation, and corruption, *Journal of Business Venturing*, 24 (5): 465-476, 2009

[3] AREND, R. J. Ethics-focused dynamic capabilities: A Small business perspective, *Small Business Economics*, 41 (1): 1-24, 2013

[4] BARON et al. Personal motives, moral disengagement, and unethical decisions by entrepreneurs: Cognitive mechanisms on the ‘Slippery Slope’, *Journal of Business Ethics*, 128 (1): 107-118, 2015

[5] BEUGELSDIJK, S. – MASELAND, R. – HOORN, A. Are Scores on Hofstede's Dimensions of National Culture Stable over Time? A Cohort Analysis, 5(3), 2015

[6] BEUGELSDIJK, S. – WELYEL, CH. Dimensions and Dynamics of National Culture: Synthesizing Hofstede With Inglehart. *Journal of Cross-Cultural Psychology*, 49 (10), 2018

[7] BOND, M. H. et al. Culture-level dimensions of social axioms and their correlates across 41 cultures. *Journal of Cross-Cultural Psychology*, 35 (5): 548-570, 2004

[8] CALABRETTA, G., DURISIN, B., OGLIENGO, M. Uncovering the intellectual structure of research in business ethics: A journey through the history, the classics, and the pillars of Journal of Business Ethics. *Journal of Business Ethics*, 104 (4), 499-524, 2011

[9] CHAN, K. C., FUNG, A., FUNG, H.-G., YAU, J. A citation analysis of business ethics research: A global perspective. *Journal of Business Ethics*, 136 (3), 557-573, 2016

[10] CHELL et al. Social entrepreneurship and business ethics: Does social equal ethical? *Journal of Business Ethics*, 133 (4): 619-625, 2016

[11] COLLEWAERT, V. - FASSIN, Y. Conflicts between entrepreneurs and investors: The impact of perceived unethical behavior, *Small Business Economics*, 40 (3): 635-649, 2013

[12] CORDEIRO, W. P. Entrepreneurial business ethics: A special case or business as usual? *International Journal of Economics and Business Research*, 3 (3): 241-252, 2011

[13] CRANE, F. G. Ethics, entrepreneurs and corporate managers: A Canadian study, *Journal of Small Business & Entrepreneurship*, 22 (3): 267-274, 2009

[14] CUMMING et al. Sustainable and ethical entrepreneurship, corporate finance and governance, and institutional reform in China, *Journal of Business Ethics*, 134 (4): 505-508, 2016

[15] DEY, P. - STEYART, C. Rethinking the space of ethics in social entrepreneurship: Power, subjectivity, and practices of freedom, *Journal of Business Ethics*, 133 (4): 627-641, 2016

[16] DROVER et al. Take the money or run? Investors' ethical reputation and entrepreneurs' willingness to partner, *Journal of Business Venturing*, 29 (6): 723-740, 2014

[17] ERINGA, K. et al. How relevant are Hofstede's dimensions for inter-cultural studies? A replication of Hofstede's research among current international business students. Research in Hospitality Management, 5 (2): 187-198, 2015

[18] FELLNHOFER et al. The current state of research on sustainable entrepreneurship*, International Journal of Business Research*, 14 (3): 163-172, 2014

[19] FRANKENA, W., K., N. Hoerster (Ed.), Ethik: Eine Analytische Einführung (6 ed.), Springer VS, Wiesbaden, 2016

[20] FULLER, T. - TIAN, Y. Social and symbolic capital and responsible entrepreneurship: An empirical investigation of SME narratives, *Journal of Business Ethics*, 67 (3): 287-304, 2006

[21] GERLACH, P. – ERIKSSON, K. Measuring Cultural Dimensions: External Validity and Internal Consistency of Hofstede's VSM 2013 Scales, *Frontiers in Psychology*, 2021

[22] GONIN et al. Managing social-business tensions: A review and research agenda for social enterprise*, Business Ethics Quarterly*, 23 (3): 407-442, 2013

[23] GULBRO, D. R. - HERBIG, P. Cultural differences encountered by firms when negotiating internationally. Industrial Management and Data Systems, 28 (2): 47-53, 1999

[24] HAMMANN et al. Values that create value: Socially responsible business practices in SMEs – Empirical evidence from German companies, *Business Ethics: A European Review*, 18 (1): 37-51, 2009

[25] HANNAFEY, F. T. Entrepreneurship and ethics: A literature review. *Journal of Business Ethics*, 46 (2): 99-110, 2003

[26] HEMINGWAY, C. A. Personal values as A catalyst for corporate social entrepreneurship. *Journal of Business Ethics*, 60 (3): 233-249, 2005

[27] HOFSTEDE, G. Culture's consequences: international differences in work-related values. CA: SAGE, Beverly Hills, 1980

[28] HOFSTEDE, G. - HOFSTEDE G. J. Kultúrák és szervezetek. Az elme szoftvere, (Cultures and Organizations: Software of the Mind). Publisher: VHE Kft, Pécs, 2008

[29] HOFSTEDE, G. - BOND, M. H. The Confucius connection: from cultural roots to economic growth, *Organizational Dynamics*, 16 (4): 5-21, 1988

[30] HOFSTEDE, G. - HOFSTEDE, G. J - MINKOV, M. Cultures and Organizations: Software of the Mind (Rev. 3rd ed.), New York: McGraw-Hill, 2010

[31] HOFSTEDE, G., VAN DEUSEN, C. A., MUELLER, C. B., CHARLES, T. A. What goals do business leaders pursue? A study in fifteen countries, *Journal of International Business Studies*, 33 (4): 785-803, 2002

[32] HOUSE, R. J. - HANGES, P. J. - JAVIDAN, M. - DORMAN, P. V. – GUPTA, V.: Culture, Leadership, and Organizations, The GLOBE Study of 62 Societies, Sage Publishing, 2004

[33] JENKINS, H. A ‘business opportunity’ model of corporate social responsibility for Small- and medium-sized enterprises, *Business Ethics: A European Review*, 18 (1): 21-36, 2009

[34] KABACK, M. M. Population-based genetic screening for reproductive counseling: The Tay-Sachs disease model, *European Journal of Pediatrics*, 159 (3): 192-195, 2000

[35] KIRKMAN, B. L., LOWE, K. B. - GIBSON, C. A quarter century of Culture's Consequences: A review of the empirical research incorporating Hofstede's cultural value framework, *Journal of International Business Studies*, 36 (3). 285-320, 2006

[36] KRAUS et al. Social entrepreneurship: An exploratory citation analysis, *Review of Managerial Science*, 8 (2): 275-292, 2014

[37] KRAUS et al. Sustainable entrepreneurship orientation: A reflection on status-quo research on factors facilitating responsible managerial practices, *Sustainability*, 10 (2): pp. 1-21, 2018

[38] KURATKO, D. F. - GOLDSBY, M.G. Corporate entrepreneurs or rogue middle managers? A framework for ethical corporate entrepreneurship, *Journal of Business Ethics*, 55 (1): 13-30, 2004

[39] LETENYEI, L. Kulturális antropológia, (Cultural Anthropology). Publisher: Typotex Kiadó, Budapest, 2012

[40] LEPOUTRE, J. - HEENE, A. Investigating the impact of firm size on small business social responsibility: A critical review, *Journal of Business Ethics*, 67 (3): 257-273, 2006

[41] MACHOVA, R., ZSIGMOND, T., LAZÁNYI, K., BENCSIK, A. Generations and Emotional Intelligence A Pilot Study, *Acta Polytechnica Hungarica,* 17 (5): 229-247, 2020

[42] MALHOTRA, N. K. Marketingkutatás, (Marketing Research). Publisher: Akadémiai Kiadó Rt., Budapest, 2005

[43] MARKMAN et al. Entrepreneurship as a platform for pursuing multiple goals: A special issue on sustainability, ethics, and entrepreneurship, *Journal of Management Studies*, 53 (5): 673-694, 2016

[44] McLEAN, J. - LEWIS, R. D. Communicating across cultures, *British Journal of Administrative Management*, 71: (30-31), 2010

[45] McVEA, J. F. - FREEMAN, R. E. A names-and-faces approach to stakeholder management: How focusing on stakeholders as individuals can bring ethics and entrepreneurial strategy together, *Journal of Management Inquiry*, 14 (1): 57-69, 2005

[46] MILENKOVIC, M. Global Advertising in a Cultural Context. Diplomica Verlag Gmbh, Hamburg, 2009

[47] MORRIS et al. The ethical context of entrepreneurship: Proposing and testing a developmental framework, *Journal of Business Ethics*, 40 (4): 331-361, 2002

[48] PEARSON et al. Predictability of physiological testing and the role of maturation in talent identification for adolescent team sports, *Journal of Science and Medicine in Sport*, 9 (4): 277-287, 2006

[49] PELLEGRINI, M. - CIAPPEI, C. Ethical judgment and radical business changes: The role of entrepreneurial perspicacity, *Journal of Business Ethics*, 128 (4): 769-788, 2015

[50] REY-MARTÍ et al. A bibliometric analysis of social entrepreneurship, *Journal of Business Research*, 69 (5): 1651-1655, 2016

[51] REYNOLDS, N., SIMINTIRAS, A. - VLACHOU, E. International business negotiations: Present knowledge and direction for future research, *International Marketing Review*, 20 (3): 236-261, 2003

[52] SACKEY et al. Working with or against the system: Ethical dilemmas for entrepreneurship in Ghana, *Journal of Developmental Entrepreneurship*, 18 (1): 1-18, 2013

[53] SIVAKUMAR, K. - NAKATA, C. The stampede toward Hofstede's framework: avoiding the sample design pit in cross-cultural research, *Journal of International Business Studies*, 32 (3): 555-574, 2001

[54] SCHWARTZ, S. H. Beyond individualism/collectivism: New cultural dimensions of values. In U. Kim, H. C., 1994

[55] SHORT, J. C., PAYNE, G. T., KETCHEN, D. J. Research on organizational configurations: Past accomplishments and future challenges, *Journal of Management*, 34 (6): 1053-1079, 2008

[56] SMITH, P. B. - DUGAN, S. - TROMPENAARS, E. National culture and managerial values: A dimensional analysis across 43 nations, *Journal of Cross-Cultural Psychology*, 27: 252-285, 1996

[57] TROMPENAARS, F. Riding the Waves of Culture. Nicholas Brealey Publishing, London, 1995

[58] VALLASTER et al. Responsible entrepreneurship: Outlining the contingencies, *International Journal of Entrepreneurial Behavior & Research,* 10.1108/IJEBR-04-2018-0206, 2018

[59] VON SCHNITZLER, A. Citizenship prepaid: Water, calculability, and techno-politics in South Africa*, *Journal of Southern African Studies*, 34 (4): 899-917, 2008

[60] WANG, R. Chinese culture and its potential influence on entrepreneurship, *International Business Research*, 5 (10): 76-90, 2012

[61] ZAHRA et al. A typology of social entrepreneurs: Motives, search processes and ethical challenges, *Journal of Business Venturing*, 24 (5): 519-532, 2009

[62] ZHANG, Z. - ARVEY, R.D. Rule breaking in adolescence and entrepreneurial status: An empirical investigation, *Journal of Business Venturing*, 24 (5): 436-447, 2009

[63] ZHU, Y. The role of Qing (positive emotions) and Li (rationality) in Chinese entrepreneurial decision making: A Confucian Ren-Yi wisdom perspective, *Journal of Business Ethics*, 126 (4): 613-630, 2015

Energy Management Optimization for Micro-grids, using a Chaotic Symbiotic Organism Search Algorithm

Bendjeghaba Omar[1], Ishak Boushaki Saida[2], Brakta Noureddine[1]

[1]Research Laboratory on the Electrification of Industrial Companies (LREEI), University M'Hamed Bougara Boumerdes, 35000, Algeria
bendjeghaba@univ-boumerdes.dz, n.brakta@univ-boumerdes.dz

[2]Department of informatics, University M'Hamed Bougara Boumerdes, 35000, Algeria, s.boushaki@univ-boumerdes.dz

Abstract: This paper presents an efficient approach, that is centered on a chaotic symbiotic organism search (CSOS) algorithm, for solving the energy management optimization (EMO) problem in Micro-grids (MG) containing diverse distributed generation resources (DGR) besides energy storage systems. The proposed approach is equipped with a chaotic map to guarantee a wider coverage of the search space and rapid time for convergence when searching solutions for the EMO problem under the various exploiting constraints. The CSOS approach is examined on a practical microgrid linked to public services. The effectiveness of CSOS is proven through a comparison of the obtained solutions, in terms of operating costs, with those of other scalable algorithms, such as, GA and PSO.

Keywords: Micro Grids; Energy Management Optimization; Distributed Generation Resources, Chaotic Symbiotic Organism search algorithm

1 Introduction

Over the past two decades, the electric power sector suffered from rapidly rising fossil fuel prices and global climate change, and researchers had to help in adopting an accepted response to save this industry from vanishing. This trend pushed the concept of clusters in this field; hence, the "Micro-grid" can be viewed as a cluster of distributed energy resources, energy storage, and local loads, managed by a smart energy management system [1] [2].

The Micro-grids (MGs) offer much superiority over traditional distribution systems, in terms of reducing energy losses due to the proximity between DGs and loads, improving reliability, it offers the ability to work in the island, to combat system failures by dividing the horizon, and added value appears as relief of transmission and distribution lines, the latter is achieved by the said energy management to reduce or by completely import the energy from the healthy grids. Whereas these benefits come with extra cost, the heavy integration of DGs will manifest as complex challenges for the MG's operation control. Therefore, for the energy management optimization (EMO) problem there is a strong need for adequate planning and location of energy sources and energy storage devices in MGs while observing satisfaction of all objectives and constraints [2].

The problem themselves is usually highly nonlinear, involving continuous and discrete variables under complex constraints, which cannot be solved by classical methods. The drawbacks from existing classical methods have shown the importance to rely on more advanced algorithms which are more adequate. Hence, the evolutionary algorithms come in terms of the solution for the drawbacks found on existing classical methods.

Recently, many evolutionary algorithms attracted intense consideration from the scientific community and formed interesting tools for solving many optimizations problem in different areas of science and industry. The main motivations towards these algorithms are due to their inherent nonlinear mapping, implementation simplicity, and powerful search capabilities [3-9].

The EMO problem in MGs consists of finding the optimal (or near-optimal) unit commitment and dispatch of available energy sources and storage devices so that certain selected criteria are met [10] [11] For this purpose, a growing number of scientific works have been developed by researchers to address and solve the problems attributed to EMO in the deterministic and probabilistic formulations. In the deterministic formulation of the EMO, it is assumed that the output variables of the DG ressource, the loads, and the market prices, are equal to their predicted values [12]. However, the uncertainty of these variables leads to a probabilistic formulation of the problem [13].

To solve the optimal power dispatch problem of interconnected MGs, while maintaining a minimum operating cost, and considering load uncertainties and generated power, Nikmehr and Ravadanegh used a PSO solution [12]. A probabilistic approach to the EOM of renewable microgrids under undefined environments is proposed in reference [13]. Hatziargyriou et al. [14] investigated the outcome of using a Microgrid Central Controller (MGCC) to ensure the coordinated operation of various DG units, storage devices, and controllable loads to avoid power losses within the local network and present the potential economic benefits. A smart energy management system, based on the matrix real-coded genetic algorithm (GA), to optimize the operation of the MG is presented in [15]. An improved PSO algorithm combined with Monte Carlo simulation is used to solve the dynamic economic dispatch of an MG system with both renewable and nonrenewable energy sources, this work has been presented in [16].

Mohan and al. in [17] proposed a stochastic weight trade-off PSO-based backward–forward sweep OPF method to obtain the online optimal schedules of DGs in MG considering renewable energy, grid power trade, and demand-side response. Chakraborty, Weiss, and Simoes proposed a linear programming algorithm to optimize the operating cost of the MG and the states of charge of the battery [18]. Tsikalakis and Hatziargyriou used centralized control of multiple MGs combined with optimizing the production of the local DGs versus power exchanges with the main distribution grid [19]. A method based on an optimal power flow and a PSO algorithm is suggested by Sortome et al to study two MGs [20]. In the paper of Mohamed et al [21], an adaptive direct mesh search algorithm is employed to minimize the cost function of MG, taking into account the cost of emissions.

An expert multi-objective Adaptive Modified Particle Swarm Optimization algorithm (AMPSO) is developed and implemented in the work of Moghaddam et al [22] to optimize the operation of a typical micro-grid with renewable energy sources accompanied by backup hybrid power sources, in this paper the problem is formulated as a multi-objective optimization problem with nonlinear constraints to simultaneously minimize the total operating cost and the net gas emission.

Mohamed and Koivo [23] applied GA for solving the EMO problem which is modeled as a nonlinear constrained multi-objective optimization, where the fitness function includes the costs of the emissions added to the start-up costs, as well as all incurring operation and maintenance costs. In the publication of Tomoiaga et al., a new heuristic approach is proposed for the energy management on stand-alone microgrids, which avoids the waste of the existing renewable potential at each time interval [24]. Nikmehr and Ravadanegh used an imperialist competitive algorithm (ICA) to solve the optimal power dispatch problem of interconnected MGs with minimum operating cost considering load uncertainties and limits of the generated power [25].

Radosavljevic´ et al. [26] presented an efficient algorithm based on PSO to tackle the EMO in an MG including different DG units and energy storage devices. Liu et al. developed an economic scheduling model of MG in grid-connected mode with the consideration of the storage battery lifetime [27].

A day-ahead optimal energy management strategy for the economic operation of industrial microgrids with high-penetration renewables under both isolated and grid-connected operation modes is well studied in the work of Han et al. [28]. The non-dominated sorting GA II is employed for optimal EMO of a grid-connected MG in the paper authored by Karuppasamypandiyan et al. [29]. To schedule power in a microgrid, the dual decomposition method was utilized in Zhang et al. [30].

The above gives a state of the art of this research field. We noticed that most of the developed approaches, by the scientific community, are based on more or less complex metaheuristics in terms of internal control parameters and random initialization. Therefore, this usually leads these approaches to a premature convergence or to get stuck in local optima.

To tackle these drawbacks, we propose a chaotic symbiotic organism search-based approach to solve the EMO problem in this paper. the performances of this approach will be evaluated and compared with some other well-known evolutionary algorithms described previously by multiple researchers [13] [22] [26].

The remainder of this paper is structured as follows: in Section 2, the constrained energy management optimization (EMO) problem is formulated. Then, the chaotic SOS (CSOS) is presented in Section 3. The case study, simulation results, and comparisons are shown in Section 4. Finally, the conclusions and future work are presented in Section 5.

2 Mathematical Formulation of the Energy Management Optimization Problem

For a practical low-voltage (LV) grid-connected MG (as shown in Figure 1) the optimization procedure depends strongly on the market policy adopted in the MG operation. In this paper, we have considered that the EMO problem is defined according to the first market policy presented in the references [14] [19] [26]. Therefore, in a typical MG, the EMO problem aims to minimize the total operating cost of the microgrid through optimal adjustment of the DG’s power generation while satisfying various system operating constraints.

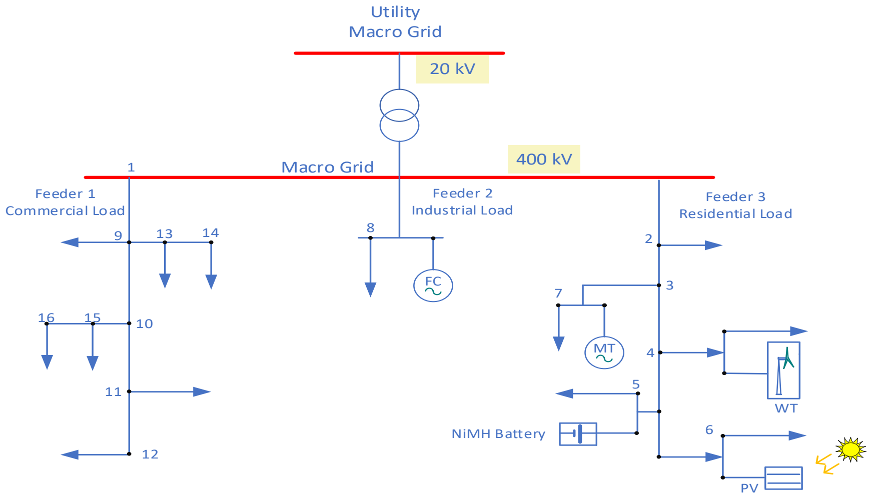


Figure 1
A typical low voltage microgrid [20]

2.1 Formulation of the EMO Objective Function

The total cost of operating the micro-grid includes the DG's offers and market prices for the exchange of electricity between the micro-grids and utilities. So, the mathematical model of such a problem concerns the minimization of the total operating cost as an objective function which can be expressed as follows:

$$Min\ F(P) = Min(\sum_{t=1}^{NT} cost^t) = \mathrm{Min}\left(\sum_{t=1}^{NT}\sum_{i=1}^{N_g}\langle B_{Gi}(P_{Gi}^t) + MP^t.P_{Grid}^t\rangle\right) \quad (1)$$

Where,

$F(P)$ is the objective function;

$P = [P^1\ P^2\ ...\ P^t\ ...\ P^{NT}]$ is a vector of candidate solution and P^t are variable-state scalar vectors including the active power of the generation and storage units within the MG, and can be described as follows:

$$P^t = \left[P_{G1}^t\ P_{G2}^t\ ...\ P_{GN_g}^t\right] \quad (2)$$

Where,

NT and is the total number of hours;

N_g is the total number of DG including storage units;

P_{Gi}^t is the real power outputs of the i^{th}DG;

$B_{Gi}(P_{Gi}^t)$is the bid of the i^{th} DG unit as a function of its active power at time t;

P_{Grid}^t is the active power which is bought (sold) from (to) the utility at time t, and MP^t is the market price of power exchange between the microgrid and the utility at time t.

2.2 Formulation of Microgrid and Unit Constraint Functions

The objective (or fitness) function formulated above is subject to the constraints due to the limits of its components: the grid power transfer limits, energy storage units' capacity and operational limits, dispatchable DGs' power limit, and all other micro-grid technical limitations and requirements.

2.2.1 Power Balance

Grid balance is guaranteed through the following considerations: for each time interval t , the combined output power of the energy storage devices of the DGs and the utility has to meet the total load demand in the micro-grid without any loss of power. Hence, the constraint of the power balance can be written as follows:

$$\sum_{i=1}^{N_g} P_{Gi}^t + P_{Grid}^t = \sum_{D=1}^{N_D} P_{L_D}^t \quad (3)$$

Where P_{L_D} is the amount of the D^{th}load level, and N_D is the total number of load levels.

2.2.2 Real Power Generation Capacity

The real power output of each unit in the microgrid, including the utility, will ensure stable operation if it is restricted by minimum and maximum power limits as follows:

$$P_{GiMin}^{t} \leq P_{Gi}^{t} \leq P_{GiMax}^{t} \tag{4}$$

$$P_{GridMin}^{t} \leq P_{Grid}^{t} \leq P_{GridMax}^{t} \tag{5}$$

Where, P_{GiMin}^{t} and $P_{GridMin}^{t}$ are the minimum active power of the i^{th} DG, and the utility at time t;

P_{GiMax}^{t} and $P_{GridMax}^{t}$are the maximum active powers of the i^{th} DG, and the utility at time t;

2.2.3 Spinning Reserve

The detected power fluctuations of renewables and load fluctuations will degrade the reliability of the system, consequently, it is necessary to adopt the spinning reserve to increase the system's reliability. The spinning reserve is met if the following inequality condition is satisfied [16] [26]:

$$\sum_{i=1}^{N_g} P_{GiMax}^{t} + P_{GridMax}^{t} \geq \sum_{D=1}^{N_D} P_{L_D}^{t} + P_{\mathrm{SSR}}^{t} \tag{6}$$

Where P_{SSR}^{t} is the scheduled spinning reserve at time t. In a microgrid, the spinning reserve constraint is considered by adding an extra value to the total power demand, which should be supplied by the DG units.

2.2.4 Energy Storage Limits

Since there are some limitations on charge and discharge rates of storage devices during each time interval, the following equation of constraints can be expressed for a typical battery as follows [22] [26]:

$$W_{ess,t} = W_{ess,t-1} + \eta_{charge}.P_{charge}.\Delta t - \frac{1}{\eta_{discharge}}.P_{discharge}.\Delta t \tag{7}$$

$$\begin{cases} W_{ess,min} \leq W_{ess,t} \leq W_{ess,max} \\ P_{charge,t} \leq P_{charge,max};\ \ P_{discharge,t} \leq P_{discharge,max} \end{cases} \tag{8}$$

Where

$W_{ess,t}$ and $W_{ess,t-1}$ are the amount of energy storage inside the battery at hour t and $(t-1)$, respectively,

$P_{charge}(P_{discharge})$ is the permitted rate of charge (discharge) during a definite period of time (t),

$\eta_{charge}(\eta_{discharge})$ is the efficiency of the battery during the charge/discharge process and $W_{ess,min}$ and $W_{ess,max}$ are the lower and upper limits on the amount of energy storage inside the battery, respectively, and $P_{charge,max}(P_{discharge,max})$ is the maximum rate of battery charge (discharge) during each time interval (Δt).

2.2.5 Calculation of the Active Power from (to) the Utility

Considering the active power from (to) the utility as a dependent variable will consequently reinforce the active power balance constraint depicted in Equation (3). Hence the value of grid power is evaluated using the following equation:

$$\left\{P_{Grid}^{t} = \sum_{D=1}^{N_D} P_{L_D}^{t} - \sum_{i=1}^{N_g} P_{Gi}^{t}\right. \tag{9}$$

The obtained P_{Grid}^{t} either satisfies the restriction defined in Equation (10) or not. Therefore, the variable $P_{Grid,lim}^{t}$ is calculated depending on P_{Grid}^{t}:

$$P_{Grid,lim}^{t} = \begin{cases} P_{Grid,min}^{t} \ if \ P_{Grid}^{t} < P_{Grid,min}^{t} \\ P_{Grid,max}^{t} \ if \ P_{Grid}^{t} > P_{Grid,max}^{t} \\ P_{Grid}^{t} \ if \ P_{Grid,min}^{t} \leq P_{Grid}^{t} \leq P_{Grid,max}^{t} \end{cases} \tag{10}$$

The control variables are said to be self-constrained, whereas the dependent variable P_{Grid}^{t}, is a relevant term in the objective function, it is considered as a quadratic penalty term. This is evaluated as a penalty factor multiplied by the square of the difference between the actual value and the limiting value of the dependent variable, which must be included in the objective function, then, all unfeasible solutions obtained during the optimization process are ignored [21]. The new extended objective function to be minimized develops to:

$$Min\ F_p(P) = Min\left(\sum_{t=1}^{NT} \sum_{i=1}^{N_g} \langle B_{Gi}(P_{Gi}^{t}) + MP^{t}.P_{Grid}^{t}\rangle + \sum_{t=1}^{NT} \alpha_p\left(P_{Grid}^{t} - P_{Grid,lim}^{t}\right)^2\right) \tag{11}$$

Where, α_p is the penalty factor.

In the above equation, the DG bids (B_{Gi}) are considered quadratic to the cost function of the units [21] [34]. They can be determined utilizing the following:

$$B_{Gi} = a_i(P_{Gi}^{t})^2 + b_i P_{Gi}^{t} + c_i \tag{12}$$

3 The Chaotic SOS Algorithm (CSOS)

The SOS algorithm is one of the most powerful optimization techniques mimicking the biological interactions between two life forms in the ecosystem, to establish a new solution for practical optimization problems. The SOS algorithm is based on three idealized phases, namely; mutualism, commensalism, and parasitism as is illustrated in Figure 2. To solve any optimization problem, the SOS iteratively uses a population of candidate solutions to explore and exploit the promising areas of the search space as presented in reference [31].

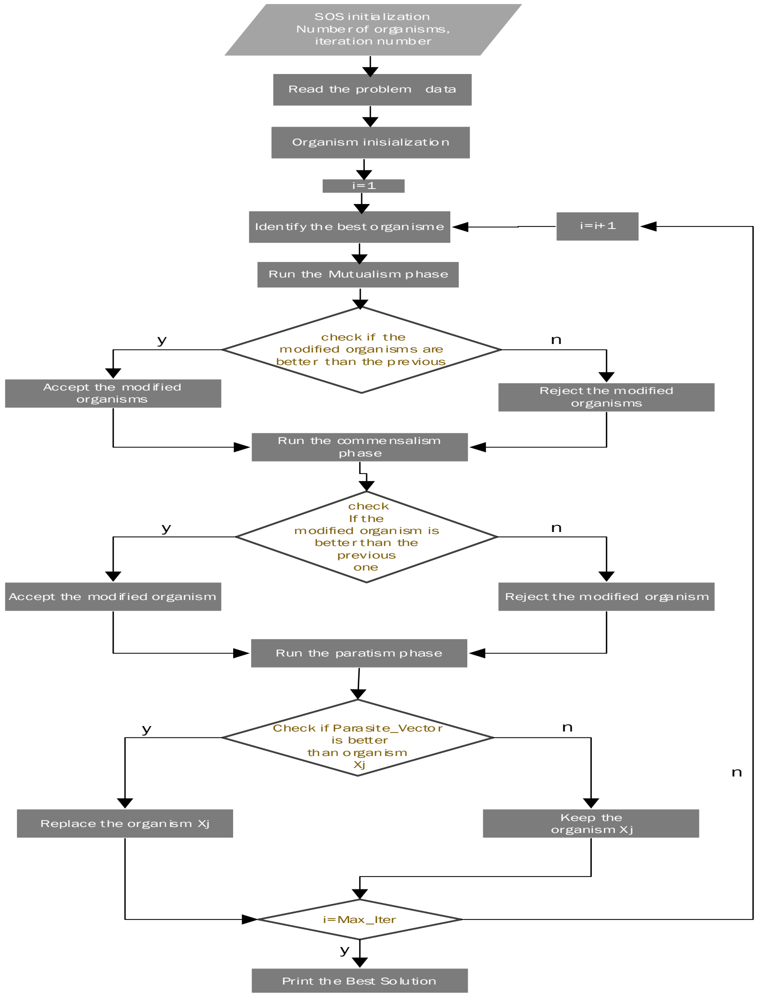

SOS initialization
Number of organisms,
iteration number
Read the problem data
Organism inisialization
i=1
Identify the best organisme
i=i+1
Run the Mutualism phase
y
check if the modified organisms are better than the previous
n
Accept the modified organisms
Reject the modified organisms
Run the commensalism phase
y
check If the modified organism is better than the previous one
n
Accept the modified organism
Reject the modified organism
Run the paratism phase
y
Check if Parasite_Vector is better than organism Xj
n
n
Replace the organism Xj
Keep the organism Xj
i=Max_Iter
y
Print the Best Solution

Figure 2

Schematic flowchart of SOS algorithm

```
1. Define input variables, objective function, and searching boundaries
% Organism size (orgsize), maximum iterations (maxiter), variables upper bound and lower bound.
2. Initialize population of organisms using the logistic map given by equation (13)
3. Identify the best organism in the initial population ($X_{best}$)
while iter<maxiter
```

```
  for i=1:orgsize
4. Mutualism Phase
        Select organisms $X_i$ and $X_j$ ($X_i \neq X_j$)
        Calculate Beneficial Factor ($BF_1$ & $BF_2$) using
        $BF_1 = 1 + round(rand(0,1))$
        $BF_2 = 1 + round(rand(0,1))$
        Calculate Mutual Vector (MV) using: $MV = ( X_i + X_j)/2$
        Generate new organisms ($X_{inew}$, $X_{jnew}$) using equations:
        $X_{inew} = X_i + rand(0,1) \times (X_{best} - MV \times BF_1)$
        $X_{jnew} = X_j + rand(0,1) \times (X_{best} - MV \times BF_2)$
        Check constraints using equations (3-10)
        Evaluate fitness value and replace predecessor if the fitness of the new organism is better
5. Commensalism Phase
        Select organism $X_j$ randomly ($X_i \neq X_j$)
        Generate new organism Xinew using
        $X_{inew} = X_i + rand(-1,1) \times (X_{best} - X_j)$
        Check constraints using equations (3-10)
        Evaluate fitness value and replace predecessor if the fitness of the new organism is better
6. Parasitism Phase
        Select organism $X_j$ randomly ($X_i \neq X_j$)
        Generate Parasite Vector (PV) by modifying $X_i$ and
        Check constraints using equations (3-10)
        Evaluate fitness value and replace $X_j$ with PV if the fitness of PV is better
  end for
7. Update best organism ($X_{best}$) of the current population
end while
8. Print the best organism ($X_{best}$) and the Best cost
```

Figure 3

Pseudo-code of CSOS algorithm

Unfortunately, the standard SOS uses a random initial population of the organism, which yields a negative impact on the efficiency of the calculation and the results. The disadvantages of this approach are its slow convergence and its tendency to be trapped in local optima due to the low diversity of the starting organism.

To improve the diversity of the initial population, many chaotic maps have been developed for the existing evolutionary algorithms [32]. In the present work, we have adopted the logistic map as an initialization strategy. This later is one of the simplest chaotic maps. Moreover, it provides initial populations that are more diversity than the random selection which ensures smarter coverage of the search space hence, it offers a lower probability of premature convergence [32] [35]. This map is given by the following equation:

$$X_{i+1} = \eta X_i(1 - X_i), 0 \leq X_0 \leq 1 \quad (13)$$

Where, X_i is the logistic chaotic value for the ith organism;

X_0 is used for generating the initial population of CSOS, $X_0 \in (0,1)$ and η is set to 4

The pseudo-code of the proposed CSOS algorithm is presented in Figure 3

4 Case Study

In this part of the work we implemented and examined the above described CSOS to find optimal global (or near-optimal) solutions of the deterministic EMO problem defined by the augmented objective function (11) and the constraints functions (3-10).

4.1 Microgrid Dataset

The system used for the case study, as shown in Figure 1, is a typical microgrid consisting of a DG unit. These DG's are a microturbine (MT), fuel cell (FC), wind turbine (WT), photovoltaic PV, and energy storage device (NiMH battery).

We assume that all DG sources deliver active power with a unity power factor. Additionally, there is a tie between the utility and the microgrid to trade energy during a day. This link will ensure power exchange as described before. For a typical day, the load demand in the microgrid consists of: a primarily residential area, an industrial feeder serving a small workshop, and a feeder for light commercial shops, with the total energy demand of 1695 kWh is the requirement for this typical day [22][26].

The test system data in terms of the supply coefficients and operating power limits of each DG unit is given in Table 1. In addition, the forecasted daily energy prices, load curve, WT, and PV power generation are presented in Figure 4

Table 1

The power limits and coefficients of bid functions of the installed DG units

Type	Min (kW)	Max (kW)	a_i (€ct/kW h2)	b_i (€ct/kW h)	c_i(€ct/ h)	Star/Shut Cost (€ct)
MT	6	30	0	0.457	0	0.96
FC	3	30	0	0.294	0	1.65

PV	0	25	0	2.584	0	0
WT	0	15	0	1.073	0	0
Battery	-30	30	0	0.38	0	0
Utility	-30	30	-	-	-	-

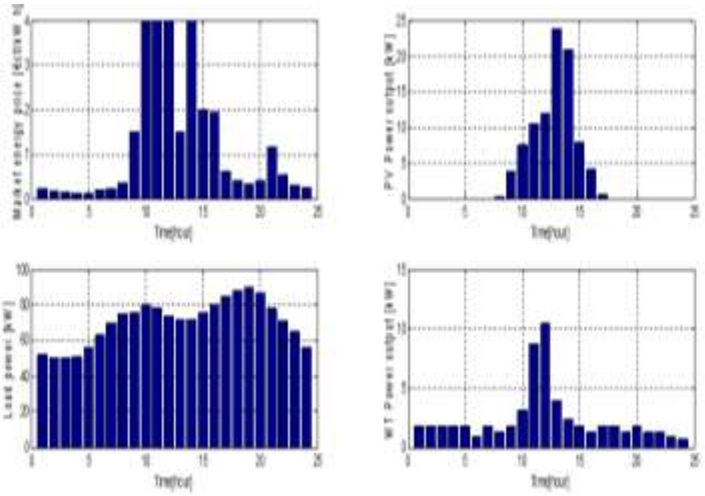


Figure 4

Daily energy market price(a), load curve(b), WT(c) and PV(d) power outputs

4.2 Simulation Results and Comparisons

To test the performance CSOS for solving EMO, initially, we have examined it for three possible operating scenarios of the considered microgrid (MG).

Then, we have conducted a comparison study of the proposed CSOS with some scalable algorithms existing in the literature. For simulations It must be noted that the developed code of CSOS for EMO is run 20 times using MATLAB R2014a software on Core i5@ 2.20 GHz, 4 GB RAM machine. Moreover, the maximum number of iterations is set at 200 with a population size of 30, and the best results are reported for each considered operating scenario.

- Scenario S1: In this scenario, we assume that both renewable energy sources (WT and PV) act at their available maximum power outputs during each hour of the day, while the remaining DGs, including MT, FC, battery, and the distribution grid (utility), operate just at their power limits yet satisfying the set constraints. All DGs above produce the electricity needed by the microgrid, however, the extra energy inside the grid is exchanged with the utility. The obtained results are presented in Figure 5.1 and Figure 5.2

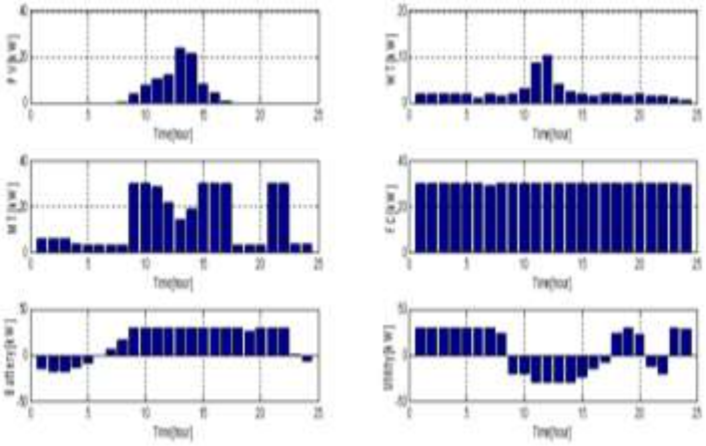


Figure 5.1

Best obtained EMO solutions using CSOS for Scenario S1

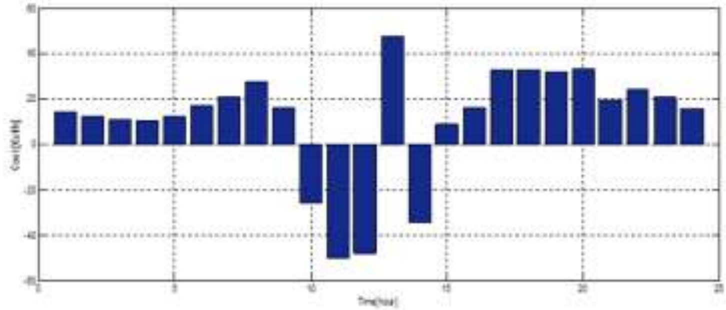


Figure 5.2

Microgrid operating cost for Scenario S1

The results of Figure 5.1 and Figure 5.2, clearly show a large part of the load is mainly supplied by the FC and the utility during the periods (1 a.m. to 8 a.m. and from 11 p.m. to midnight); obviously, this is justified by the supremacy of these 2 unites' offers compared to those of other units during the same period examined. We also notice that the excess energy is exported from the MG to the utility during the period where the prices market are much higher (from 9 a.m. to 5 p.m. and from 9 p.m. to 10 p.m.). However, the battery is charged only during the hours of the day when market prices are low.

- Scenario S2: In this scenario, we assume all the DGs and the utility operating just at their power limits yet satisfying the set constraints. The obtained results are presented in Figure 6.1 and Figure 6.2

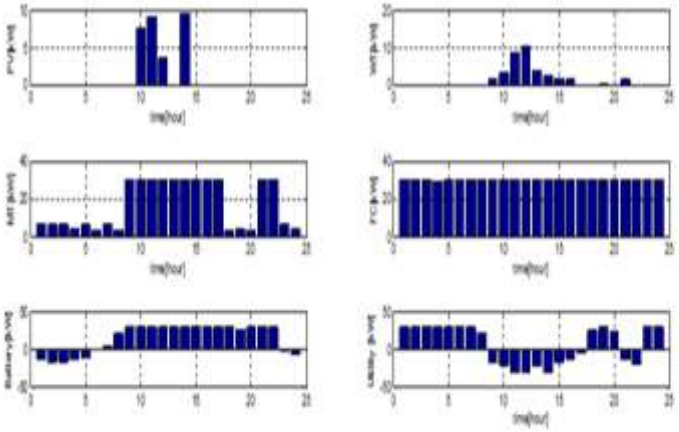


Figure 6.1
Best obtained EMO solutions using CSOS for Scenario S2

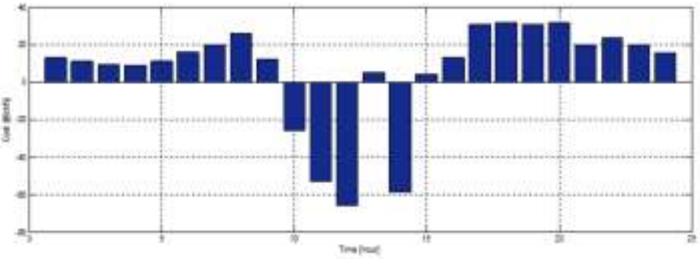


Figure 6.2
Microgrid operating cost for Scenario S2

In the second scenario, Figure 6.1 and Figure 6.2 show that the operating cost of the microgrid decreases considerably (153.98507 [€ct/day]) compared to the first scenario (268.44724 [€ct/day]). This is largely due to the significantly lower participation of WT and PV (they have much higher bids than the other DGs). The output power of FC has a maximum value throughout the day, while the MT offers change depending on the market prices. The battery charging still happens during periods of low market prices, and the extra energy is exchanged from the utility to the microgrid during the same periods.

- Scenario S3: In this scenario, we suppose that the utility behaves as an unconstrained unit and exchanges energy with the microgrid without any limitations, while the rest of the DGs act as described in the second scenario (S2). The found solutions are presented in Figure 7.1 and Figure 7.2

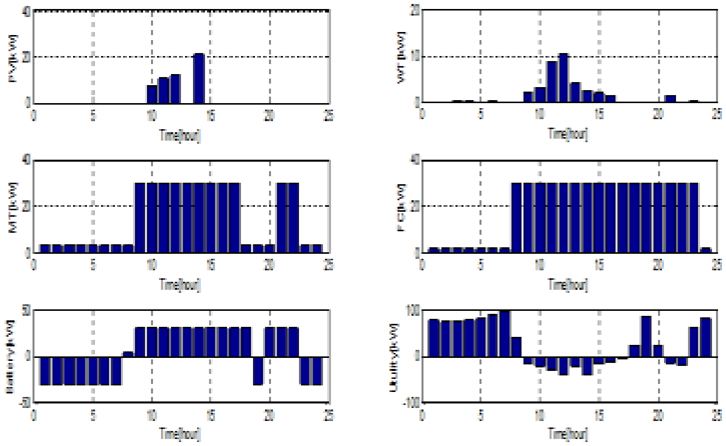


Figure 7.1

Best obtained EMO solutions using CSOS for Scenario S3

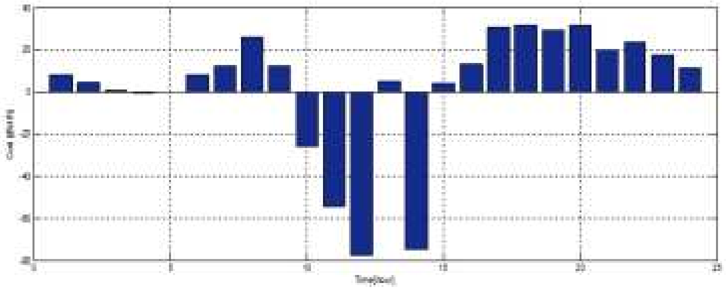


Figure 7.2

Microgrid operating cost for Scenario S3

In this scenario, the PV and the WT will start offering when a shortage of electricity production occurs inside the microgrid or when it is necessary to export more energy to the utility. Similarly, the FC, MT, and battery adjust their generation levels according to the load levels at each hour of the day.

In this situation of unlimited electricity exchange, the obtained results show a clear reduction of the microgrid operating costs (59.69627 €ct/day) compared to the first and second scenarios.

The convergent characteristics for all the scenarios considered are presented in Figure 8. So, this figure allows us to state that the CSOS exhibits a characteristic of rapid convergence and it can reach the minimum cost after a few iterations.

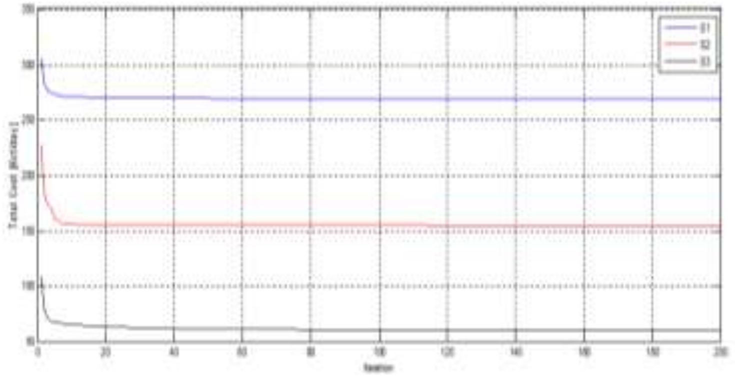


Figure 8

Convergence characteristics of CSOS for the operation scenarios

4.3 CSOS Comparisons with other Evolutionary Algorithms

While using the same test system, control variable limits, and constraints, the obtained results for EMO after deploying the proposed CSOS approach are compared to some other well-known evolutionary algorithms deployment results briefly described in the introduction with their corresponding references. The comparison results are presented in Table 2.

The comparison of results when sorted from best and worst cost, clearly shows that the suggested approach reveals a better performance with the other considered evolutionary algorithm and the standard SOS for all considered scenarios.

Concerning the execution speed, the total execution time of the CSOS is 44.02 seconds. Although it is very hard to compare execution time with other research in literature without enough information about their execution times, it can still be noticed that CSOS may be a successful candidate when it comes to execution speed as it only used 120 iterations to achieve better results in comparison with 1500 iterations for PSO, AMPSO-L, GSA, and GSA as indicated in [13] [22]. Moreover, the standard deviation confirms well another advantage of the CSOS in the optimization process. Hence, we are taking the number of iterations as a measure to say that this a better candidate.

Table 2

Best obtained solutions for EMO using CSOS over other metaheuristics

Scenarios	Method	Study reference	STD	Best cost	Worst cost
S1	GA	[22]	13.4421	277.7444	304.5889
	PSO	[22]	10.1821	277.3237	303.3791
	AMPSO-T	[22]	0.321	274.5507	275.0905
	AMPSO-L	[22]	0.0921	274.4317	274.7318

	GSA	[13]	2.9283	275.5369	282.1743
	SGSA	[13]	0	269.76	269.76
	PSO	[26]	0	269.75999	269.75999
	SOS		**0**	**269.75977**	**269.75977**
	CSOS		**0**	**268.44724**	**268.44724**
S2	GA	[22]	24.5125	162.9469	198.5314
	PSO	[22]	12.6034	162.0038	180.2282
	AMPSO-T	[22]	0.3427	159.9244	160.4091
	AMPSO-L	[22]	0.0963	159.3628	159.6813
	PSO	[26]	0	155.01333	155.01333
	SOS		**0**	**155.01324**	**155.01324**
	CSOS		**0**	**153.98507**	**153.98507**
S3	GA	[22]	13.4005	91.3293	127.7625
	PSO	[22]	10.8689	90.7629	112.8628
	AMPSO-T	[22]	0.4457	89.9917	90.6221
	AMPSO-L	[22]	0.0921	89.972	90.0431
	PSO	[26]	0	68.17626	68.17626
	SOS		**0**	**68.17626**	**68.17626**
	CSOS		**0**	**59.69627**	**59.69627**

Conclusions

This paper introduces a chaotic symbiotic search algorithm, for solving energy optimization management problems, under different operational conditions, for microgrids. In this regard, the suggested CSOS approach includes a chaotic map to improve the approach's search capabilities and ensure better convergence in terms of finding the optimal results and time for convergence.

The effectiveness of the suggested approach was examined under three different operating conditions (scenarios) and compared with other well-known population-based evolutionary algorithms, that were described previously, by other researchers. The results achieved using the CSOS are very interesting, in terms of performance and rapid convergence. Furthermore, the recommended approach doesn’t require any internal control parameter. This study was limited to the deterministic case of EMO, thus, we plan to extend it, to the probabilistic case, when the outputs of DGs and the load demand, are both variable over time and difficult to predict.

References

[1] Bahramirad S. and Daneshi H., ‘Optimal sizing of smart grid storage management system in a microgrid’. in Innovative Smart Grid Technologies (ISGT), IEEE PES, Washington, DC, USA, 2012:1-7

[2] Fu Q., Montoya L.F., Solanki A., Nasiri A., Bhavaraju V., Abdallah T., and Yu D.C., 'Microgrid generation capacity design with renewables and energy storage addressing power quality and surety'. IEEE Transactions on Smart Grid 2012; 3(4): 2019-2027

[3] Ishak Boushaki S., Kamel N., Bendjeghaba O. 'A new quantum chaotic cuckoo search algorithm for data clustering'. Expert Systems with Applications 2018a; 96 (15): 358-372

[4] Ishak Boushaki S., Kamel N., Bendjeghaba O. 'Biomedical Document Clustering Based on Accelerated Symbiotic Organisms Search Algorithm'. International Journal of Swarm Intelligence Research. 2021; 12(4): 169-185

[5] Roman R. C., Precup R. E., David R. C. 'Second Order Intelligent Proportional-Integral Fuzzy Control of Twin Rotor Aerodynamic Systems'. Procedia Computer Science. 2018; (139): 372-380

[6] Moattari M. and Moradi M. H. 'Conflict Monitoring Optimization Heuristic Inspired by Brain Fear and Conflict Systems'. International Journal of Artificial Intelligence. 2020; 18(1): 45-62

[7] Ján Č and Ján J. 'Choosing the Optimal Production Strategy by Multi-Objective Optimization Methods'. Acta Polytechnica Hungarica. 2020; 17(5): 7-26

[8] Henry Z., Niriaska P., Wilfredo A., Joyne C., 'A hybrid swarm algorithm for collective construction of 3D structures'. International Journal of Artificial Intelligence. 2020; 18(1); 1-18

[9] Precup R. E., Hedrea E. L., Roman R. C., Petriu E. M., Szedlak S., Alexandra I., Bojan D., and Claudia A. 'Experiment-based approach to teach optimization techniques'. IEEE Transactions on Education. 2021; 64 (2): 88-94

[10] Bollen M., Zhong J., Lin Y., 'Performance indices and objectives for microgrids, Proc. of 20th International Conference on Electricity Distribution. Prague, (June 8-11) (2009) Paper 0607

[11] Kanchev H., Lu D., Colas F., Lazarov V., Francois B., 'Energy Management and Operational Planning of a Microgrid With a PV-Based Active Generator for Smart Grid Applications', IEEE Trans. on Industrial Electronics 2011; 58: 4583-4592

[12] Nikmehr N., Ravadanegh S. N. 'Optimal operation of distributed generations in micro-grids under uncertainties in load and renewable power generation using heuristic algorithm'. IET Renewable Power Generation. 2015; 9(8):982-990

[13] Niknam T., Golestaneh F., Malekpour A. 'Probabilistic energy and operation management of a microgrid containing wind/photovoltaic/fuel cell

generation and energy storage devices based on point-estimate method and self-adaptive gravitational search algorithm'. Energy. 2012; 43:427-437

[14] Hatziargyriou N. D., Anastasiadis A. G., Tsikalakis A. G., Vasiljevska J. 'Quantification of economic, environmental and operational benefits due to significant penetration of microgrids in a typical LV and MV Greek network'. European Transaction on Electrical Power. 2011; 21(2):1217-1237

[15] Chen C., Duan S., Cai T., Liu B., Hu G. 'Smart energy management system for optimal microgrid economic operation'. IET Renewable Power Generation. 2011; 5:258-267

[16] Wu H., Liu X., Ding M. 'Dynamic economic dispatch of a microgrid: mathematical models and solution algorithm'. International Journal of Electrical Power & Energy Systems. 2014; 63:336-346

[17] Mohan V., Singh J. G., Ongsakul W., Suresh M. P. R. 'Performance enhancement of online energy scheduling in a radial utility distribution microgrid'. International Journal of Electrical Power & Energy Systems. 2016; 79:98-107

[18] Chakraborty S., Weiss M. D., Simoes M. G. 'Distributed intelligent energy management system for a single-phase high-frequency AC microgrid'. IEEE Transactions on Industrial Electronics. 2007; 54:97-109

[19] Tsikalakis A. G., Hatziargyriou N. D. 'Centralized control for optimizing microgrids operation. IEEE Transactions on Energy Conversion. 2008; 23(1):241-248

[20] Sortome E., El-Sharkawi M. A. 'Optimal power flow for a system of microgrids with controllable loads and battery storage. IEEE/PES Power Systems Conference and Exposition; Seattle, WA, USA, Mar 2009. IEEE; 2009: 1-5

[21] Mohamed F. A., Koivo H. N. 'System modeling and online optimal management of MicroGrid using Mesh Adaptive Direct Search'. International Journal of Electrical Power & Energy Systems. 2010; 32:398-407

[22] Moghaddam A. A., Seifi A., Niknam T., Pahlavani M. R. A. 'Multi-objective operation management of a renewable MG (micro-grid) with back-up micro-turbine/fuel cell/battery hybrid power source'. Energy. 2011;36 (18): 6490-6507

[23] Mohamed F. A., Koivo H. N. 'Online management genetic algorithms of microgrid for residential application'. Energy Conversion and Management. 2012; 64:562-568

[24] Tomoiaga B., Chindric M., Sumper A., Marzband M. 'The optimization of microgrids operation through a heuristic energy management algorithm'. Advanced Engineering Forum. 2013: 8:185-194

[25] Nikmehr N., Ravadanegh S. N. ‘A study on optimal power sharing in interconnected microgrids under uncertainty. International Transactions on Electrical Energy Systems. 2016; 26(1):208-232

[26] Radosavljev J., Jevti M., Klimenta D., “Energy and operation management of a microgrid using particle swarm optimization”. Engineering Optimization. 2015; 47(6):1-20

[27] Liu C., Wang X., Wu X., Guo J. ‘Economic scheduling model of microgrid considering the lifetime of batteries’. IET Generation, Transmission & Distribution. 2017; 11(3):759-767

[28] Han L., Abinet T. E., Jianhua Z., and Dehua Z., ‘Optimal energy management for industrial microgrids with high-penetration renewables’.Protection and Control of Modern Power Systems. 2017: 2-12

[29] Karuppasamypandiyan M., Jeyanthy P. A., Devaraj D., and Idhaya Selvi V. A., "An Efficient Non-dominated sorting Genetic algorithm II (NSGA II) for Optimal Operation of Microgrid," 2019 IEEE International Conference on Clean Energy and Energy Efficient Electronics Circuit for Sustainable Development (INCCES), Krishnankoil, India, 2019: 1-6

[30] Zhang Y., Gatsis N., Giannakis G. B. ‘Robust energy management for microgrids with high-penetration renewables’. IEEE Transactions on Sustainable Energy. 2013; 4(4):944-53

[31] Min-Yuan C., Doddy P. ‘Symbiotic Organisms Search: A new metaheuristic optimization algorithm’. Computers & Structures,2014, 139: 98-112

[32] Kaveh, A., Mahdipour Moghanni, R. & Javadi, S. M. Optimum design of large steel skeletal structures using chaotic firefly optimization algorithm based on the Gaussian map’. Structural and Multidisciplinary Optimization, 2019, 60(3):879-894

[33] Chou, JS., Ngo, NT. ‘Modified firefly algorithm for multidimensional optimization in structural design problems. Structural and Multidisciplinary Optimization, 2017, 55(6):2013-2028

[34] Atwa Y. M., El-Saadany E. F., Salama M. M. A., Seethapathy R., Assam M., Conti S. ‘Adequacy evaluation of distribution system including wind/solar DG during different modes of operation. IEEE Transactions on Power Systems. 2011; 26(4):1945-1952

[35] Saremi, S., Mirjalili S. and Lewis, A. ‘Biogeography-based optimization with chaos’. Neural Comput & Applic. 2014; 25: 1077-1097

Knowledge Management Challenges during COVID-19

Andrea Bencsik

J. Selye University Bratislavska cesta 3322. 94501 Komarno, Slovakia
email: bencsik.andrea@gtk.uni-pannon.hu; bencsika@ujs.sk

Abstract: The efficiency of organizational processes largely depends on the quality of Knowledge Management. In the crisis situation, caused by the Coronavirus, its significance is especially apparent. Our research sought to reveal what knowledge management problems have emerged, due to the spread of the Coronavirus and what difficulties need to be coped with, in organizations. During the research, in three small groups within an online workshop, a group of experts worked with the "Be-novative" program, which uses design thinking, to connect the innovation process with the power of gamification and crowdsourcing. Using the program, ideas were formulated through joined-up thinking, evaluated online and further developed. Out of the 141 problems/ideas raised, based on the community's evaluation, three complex solution possibilities were developed, which combine several ideas under comprehensive titles. The developed proposals were published on the website of the professional organization, thus, supporting the successful functioning of knowledge management programs.

Keywords: Be-novative; Covid-19; knowledge management; online

1 Introduction

We live among constantly changing factors; the tools and developments are ephemeral, becoming obsolete in months. The compulsion of constant renewal does not only affect large companies. The question is who can constantly create new ideas, technologies and present something new from time to time. The efficiency and effectiveness of organizational processes depend on the quality of knowledge management. In difficult times, such as, increased fluctuation, the departure of the aging labor force, recurring failures, crisis situations, etc., its significance is especially apparent [1].

The emergence and rapid spread of the new coronavirus turned the world upside down. Societies, economies, organizations, and individuals face the invisible enemy, seeking a way to solve hitherto unknown problems. Relying on formerly acquired knowledge, knowledge sharing and joined-up thinking can help with the search [2] [3]. In recent weeks, many have experienced that professionals are facing new Knowledge Management (KM) difficulties. Every day, there are new questions on social media: How do you solve this now? Since many are affected by the topic, it is worth thinking over how the scattered existing knowledge can be enriched, what tools can reveal individual knowledge and turn it into a shared resource. What solutions are worth developing that can be used even after the crisis? Joined-up thinking was started by these thoughts. The research question is: What knowledge management problems have emerged due to the situation caused by the spread of the coronavirus, what difficulties need to be coped with in organizations? To answer the question, before the practical examination, the lessons of past crisis situations are briefly recalled, and the most important solutions of knowledge management that fit the present situation are highlighted.

2 Theoretical Overview

Humanity has lived through several crisis situations, including natural disasters, wars and various economic catastrophes. The crisis caused by the coronavirus – like any other crisis – requires comprehensive problem treatment. A stable state of the relatively predictable past could be enjoyed during decades. Knowledge acquisition and sharing were unhindered for those who considered it important and used its tools [4]. The present is chaos and disintegration that are uncontrolled and the future…? For its estimation, experience from the peaceful past can be used limitedly, so more creativity, joined-up thinking, and knowledge sharing are required to master the present chaos and then focus on the future. In the further part of the study, the organizational level is decisive in terms of thinking, stating that the operational framework of organizations is influenced by a higher level medium, the economic possibilities, and the social status of individual countries.

2.1. Lessons from the Past

Crises are well-known in history. (The Global Economic crises starting in the USA in 1857 and 1919, 1873, the worst crisis in the 19th Century, the oil crises of the 1970s, and the savings and credit crisis of 1989, in 1929-33 Great Depression. In the "housing bubble" of 2007-2008, trading with risky loans and overconsumption caused a critical situation) [5]. What can organizations learn from these events? What possibilities do they have that can mean survival and long-term success of the organization after waning problems?

The most important task for the leaders of organizations is to outline the future of their organizations, which requires immediate decisions and reasonable resource distribution. The thinking of Henry Ford, a great personality of the above-mentioned historical crises, is still part of the university curriculum. As a result of his actions during the Great Depression, he not only survived but left a successful company for his successors [6]. The other, though not global, crisis (still part of the curriculum in several countries) is the Shackleton model, an excellent example of managerial solutions used during the famous South Pole expedition [7]. Instead of continuing crisis situation successes, let the conclusion be summarized.

- Vision, goal formulation
- Resource alignment/distribution
- Selecting/retaining talented people
- Meaningful tasks for everyone
- Appreciation (material, physical, human)
- Cooperation
- Team Spirit
- Knowledge transfer, sharing
- Getting the best out of everyone
- Setting examples
- Supporting people, mentoring

In the next section, the lessons that can be used for knowledge management are reviewed.

2.2. Necessary Elements of KM

Knowledge Management (KM) has long been essential for businesses to harness organizational expertise to make informed decisions and achieve optimal efficiency. The pandemic has increased the importance of KM, and, as in many other areas of personal and professional life, the changes have also impacted the practice of knowledge management [8].

Knowledge is a strategic resource that helps decision-makers manage the pandemic and mitigate its health and socio-economic impacts. Unlike other disasters, pandemics have a long lead time. Despite the dominant role of knowledge management in pandemics, the literature on the subject is scarce, with only health journals addressing it and knowledge management scholars being only tangentially concerned with it [9].

Looking at the key lessons, most of them can serve as a roadmap today. The factors that serve the success of companies and the development of order over chaos are in line with the characteristics and applied tools of knowledge management. Setting goals, hiring talented people, and leveraging and sharing their knowledge, as well as collaboration, team spirit, and mentoring are all features of the KM system. It is reasonable to assume that the development of a knowledge management system is a criterion of organizational functioning that is key for the development of a successful future [10].

Before the emergence of COVID-19, very few organizations used knowledge management tools well. KM is indeed not easy; it can be costly, it requires a commitment that many organizations lack the capacity to make, and the payback often takes longer than expected. Although organizational knowledge is needed even more in this crisis than ever [11].

The outbreak of the COVID-19 pandemic and the resulting social distancing requirements have led to significant disruptions in the world of work. The results of the constrained and extensive practice of working from home are still largely unexplored. The increase in physical distances has emphasized the links between people. Information flow increasingly relies on digital tools and virtual experiences to communicate and to maintain the work done. This situation highlights opportunities that have not been considered so far. The online application of innovative KM solutions, collaborative programs opens up new opportunities [12]. The former face-to-face solutions are replaced by virtual space. Although the new solutions do not require physical presence, programs have been created that support the generation and further development of common ideas by providing conditions of virtuality. Although it is not personally possible to collect, evaluate and reflect on the generated knowledge and new ideas, with the virtual tools of 'gaming' solutions better results can be achieved in a time-saving way. [13] [14].

Choosing from the models that can be applied in practice, it is now worth focusing on what can ensure the collection, further development of scattered (often individual) knowledge, and its transformation into a common resource [15-18]. Although a number of studies have previously addressed tools to support knowledge sharing in virtual organizations [19] this situation calls for further solutions. The study reports on one of the methodologies tested in practice. The next section presents the logic of applying the method and the results of the practical research.

3 New, Innovative Methodology

The authors wanted to gather challenges and difficulties of KM to answer the research questions. To do so, a globally innovative domestic application (with numerous international awards) was used [20]. Be-novative is a cloud-based SaaS that exploits individual and collective creativity to help employees in a given company – even looser communities of up to several thousand people, – formulate their ideas through joined-up thinking, evaluate and further develop them together. It uses design thinking to connect the innovation process with the power of gamification and crowdsourcing. The Be-novative program provides an opportunity for participants to develop their companies and support innovation processes by implementing ripe innovations [20].

The Be-novative program overcomes temporal and spatial limitations and will be an excellent opportunity for the organization to gather the best ideas and knowledge. The innovative knowledge-sharing web interface quickly provides useful results even for many participants. The participants share their ideas and development suggestions in a 25 to 30-minute creative "game", then evaluate them to find new ways to develop products, organizations, companies, knowledge. With a virtual brainstorming session, the participants can find solutions to global or even to everyday problems, using the power of creativity and community [20]. Figure 1 shows a sketch of the workflow.

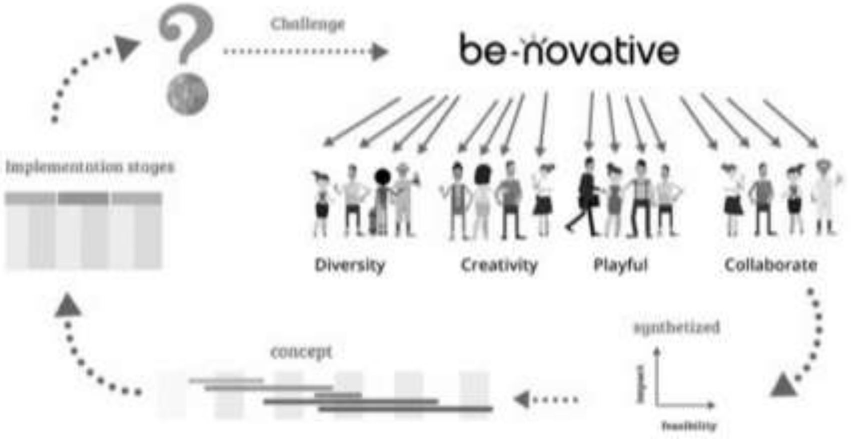
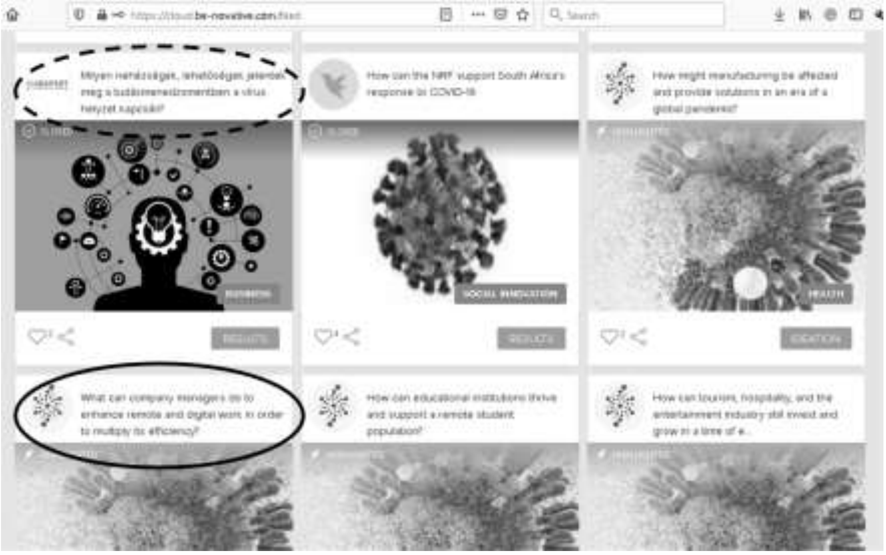


Figure 1
Workflow outline of the Be-novative program (www.be-novative.com)

In order to follow the logic, the screen views are presented via an example 'challenge', (since the research was not deducted in English, its actual screen views cannot be presented). The detailed description of our own research takes place in the Results section. The web application covers a creative process: first, various topics and issues are on the opening interface. Users can select an interesting topic or generate their own problem. In the screen view, the sample example is indicated by a solid line, and the own topic (What difficulties and opportunities have arisen in knowledge management in connection with the virus situation?) by a dashed line (Figure 2).



Milyen nehézségek, lehetőségek jelentek meg a tudásmenedzsmentben a vírus helyzet kapcsán?

How can the NRF support South Africa's response to COVID-19

How might manufacturing be affected and provide solutions in an era of a global pandemic?

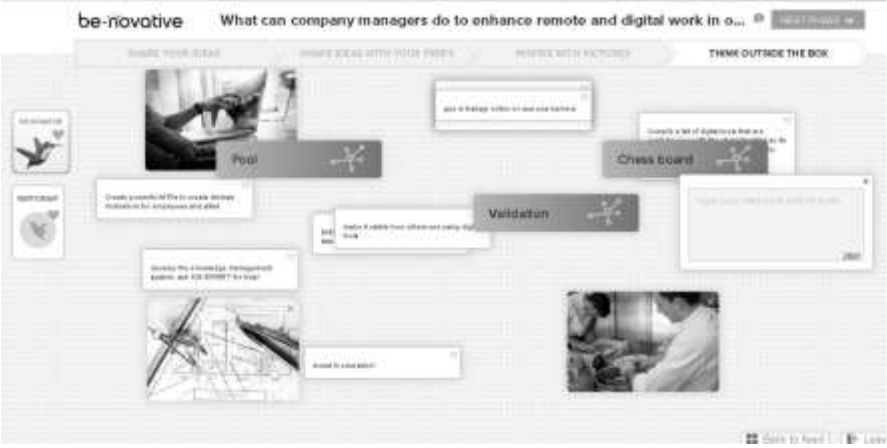
What can company managers do to enhance remote and digital work in order to multiply its efficiency?

How can educational institutions thrive and support a remote student population?

How can tourism, hospitality, and the entertainment industry still invest and grow in a time of e...

Figure 2
Selecting/generating a ‘challenge’/topic (www.be-novative.com)

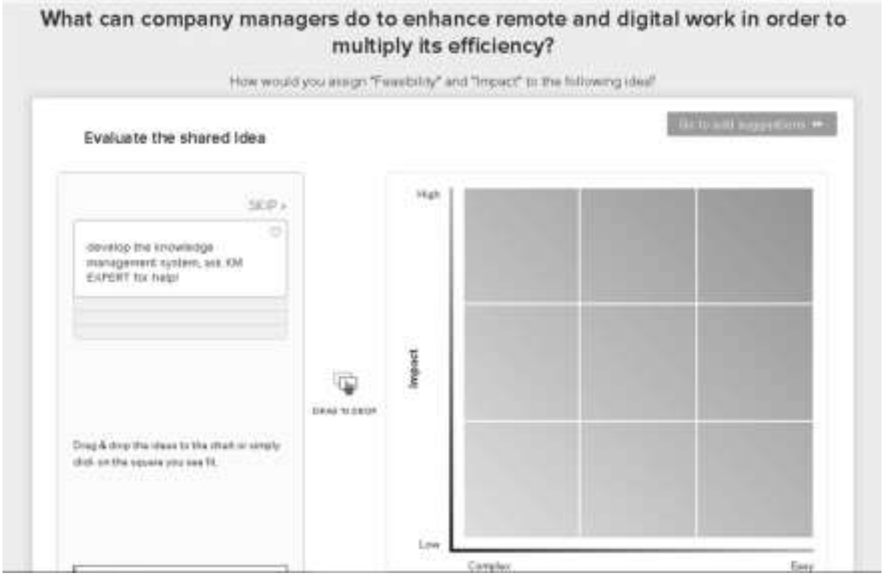
By entering a ‘challenge’, users can share their ideas about a given problem in virtual notepads, or get inspired by others’ thoughts. The program allows the users to insert images, and random words and questions appear on the screen to support the birth of completely new, breakthrough ideas, and this increases ingenuity (Figure 3).



What can company managers do to enhance remote and digital work in o...

Figure 3
Virtual notepad (www.be-novative.com)

Anonymous brainstorming lasts for fifteen minutes, and then participants evaluate suggestions in ten minutes, without knowing which idea belongs to whom (Fig. 4).



What can company managers do to enhance remote and digital work in order to multiply its efficiency?

How would you assign "Feasibility" and "Impact" to the following idea?

Evaluate the shared Idea

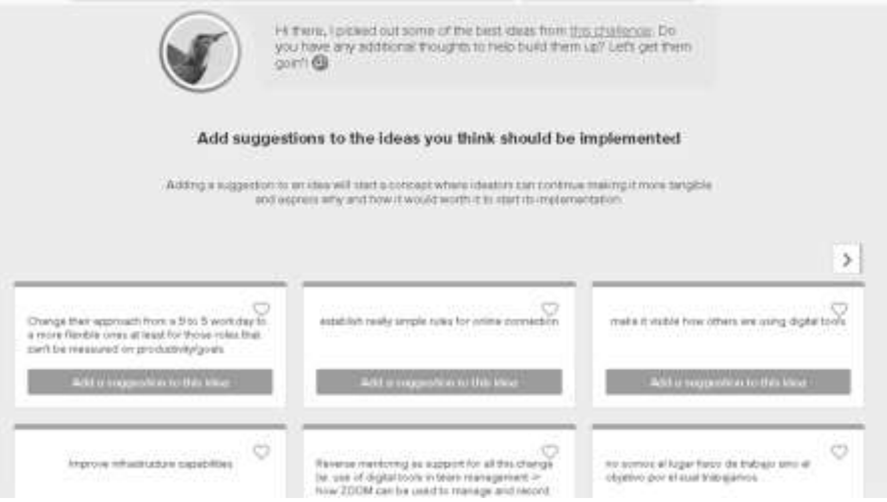
develop the knowledge management system, ask KM EXPERT for help!

Drag & drop the ideas to the chart or simply click on the square you see fit.

Figure 4

Evaluation of ideas (www.be-novative.com)

The results are averaged and arranged in a graph by the computer to show what solutions have been made for the problem (Figure 5).



Hi there, I picked out some of the best ideas from this challenge. Do you have any additional thoughts to help build them up? Let's get them goin'!

Add suggestions to the ideas you think should be implemented

Adding a suggestion to an idea will start a concept where ideators can continue making it more tangible and express why and how it would worth it to start its implementation.

Change their approach from a 9 to 5 work day to a more flexible ones at least for those roles that can't be measured on productivity/goals

establish really simple rules for online connection

make it visible how others are using digital tools

Improve infrastructure capabilities

Reverse mentoring as support for all this change (ie. use of digital tools in team management -> how ZOOM can be used to manage and record

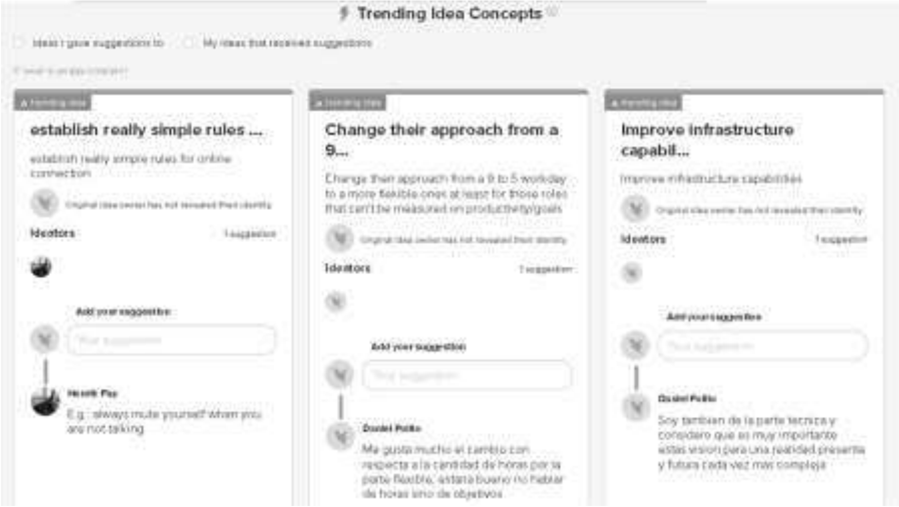
no somos el lugar fisico de trabajo sino el objetivo por el cual trabajamos

Add a suggestion to this idea
Add a suggestion to this idea
Add a suggestion to this idea

Figure 5

Solution proposals (www.be-novative.com)

Then, it is possible to further develop individual ideas and formulate more detailed proposals. (Figure 6)



Trending Idea Concepts

□ Ideas I gave suggestions to □ My ideas that received suggestions

[illegible]

Trending idea

establish really simple rules ...

establish really simple rules for online connection

Original idea owner has not revealed their identity

Ideators 1 suggestion

Add your suggestion:

Your suggestion

Henrik Fay

E.g.: always mute yourself when you are not talking

Trending idea

Change their approach from a 9...

Change their approach from a 9 to 5 workday to a more flexible ones at least for those roles that can't be measured on productivity/goals

Original idea owner has not revealed their identity

Ideators 1 suggestion

Add your suggestion

Your suggestion

Daniel Pollo

Me gusta mucho el cambio con respecta a la cantidad de horas por la parte flexible, estaria bueno no hablar de horas sino de objetivos

Trending idea

Improve infrastructure capabil...

Improve infrastructure capabilities

Original idea owner has not revealed their identity

Ideators 1 suggestion

Add your suggestion

Your suggestion

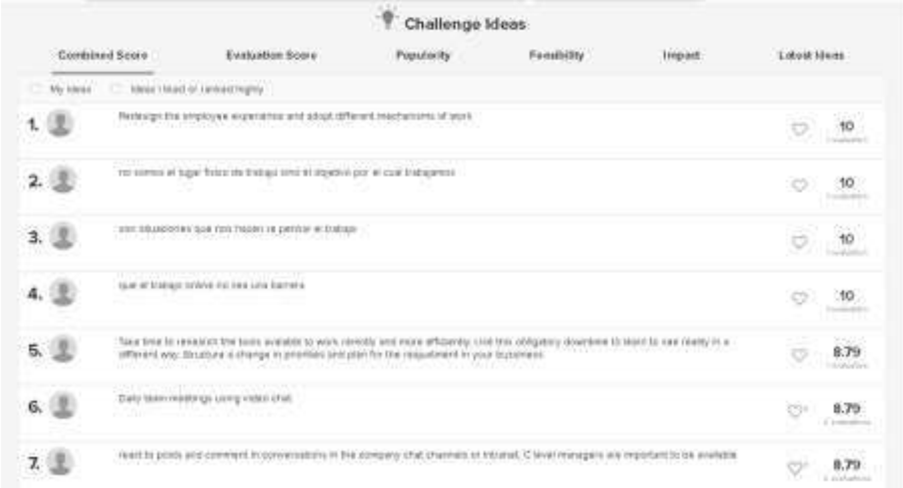
Daniel Pollo

Soy tambien de la parte tecnica y considero que es muy importante estas vision para una realidad presente y futura cada vez mas compleja

Figure 6

Further development of proposals (www.be-novative.com)

Participants can 'like' the gathered ideas, then the results summarized by the program can be queried according to various aspects (Figure 7). The owners of the ideas to be implemented can reveal themselves at the end of the process. Figure 8 presents the statistical results summarized by the software.



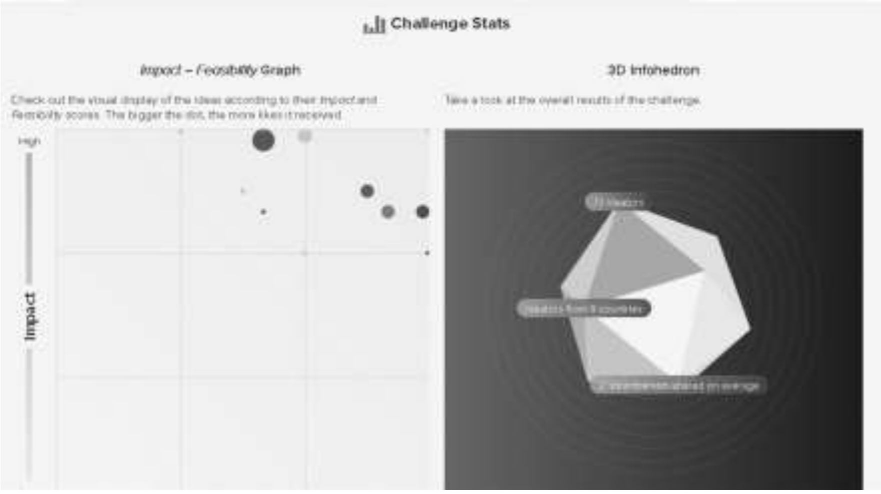
Challenge Ideas

□ My ideas □ Ideas I liked or ranked highly

1. Redesign the employee experience and adopt different mechanisms of work 10
2. no somos el lugar fisico de trabajo sino el objetivo por el cual trabajamos 10
3. son situaciones que nos hacen re pensar el trabajo 10
4. que el trabajo online no sea una barrera 10
5. Take time to research the tools available to work remotely and more efficiently. Use this obligatory downtime to start to see reality in a different way. Structure a change in priorities and plan for the requirement in your business. 8.79
6. Daily team meetings using video chat 8.79
7. react to posts and comment in conversations in the company chat channels or intranet. C level managers are important to be available 8.79

Figure 7

Ranking of ideas (www.be-novative.com)



Challenge Stats

Impact – Feasibility Graph

Check out the visual display of the ideas according to their Impact and Feasibility scores. The bigger the dot, the more likes it received.

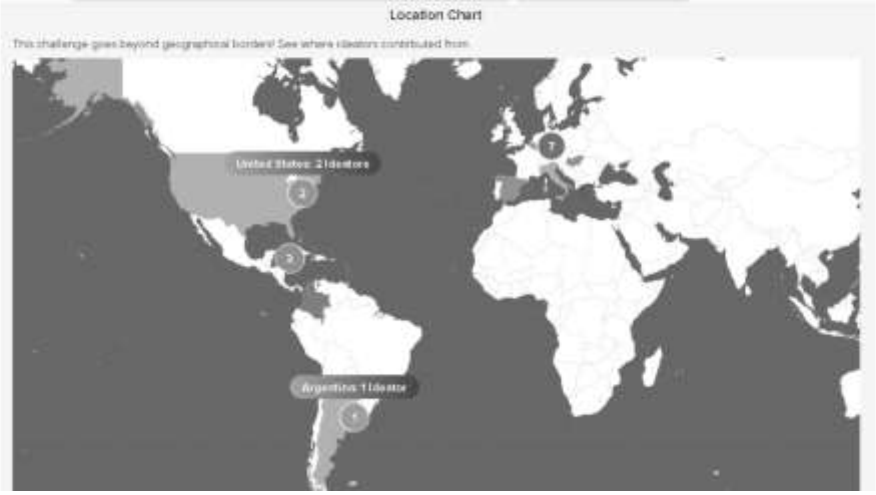
3D Infohedron

Take a look at the overall results of the challenge.

Figure 8

Statistical results of ideas (www.be-novative.com)

The brainstorming session is time-bound at a time, but within a given interval (a freely chosen daily or weekly time frame), anyone, anytime, anywhere can, users do not have to be online at the same time. As long as the interface is open, additional ideas can be entered, or existing ones evaluated, further developed, with helpful suggestions for the work. This is especially beneficial for multinational companies, given the time lag between continents. As Figure 9 shows, employees can join from multiple parts of the world.



Location Chart

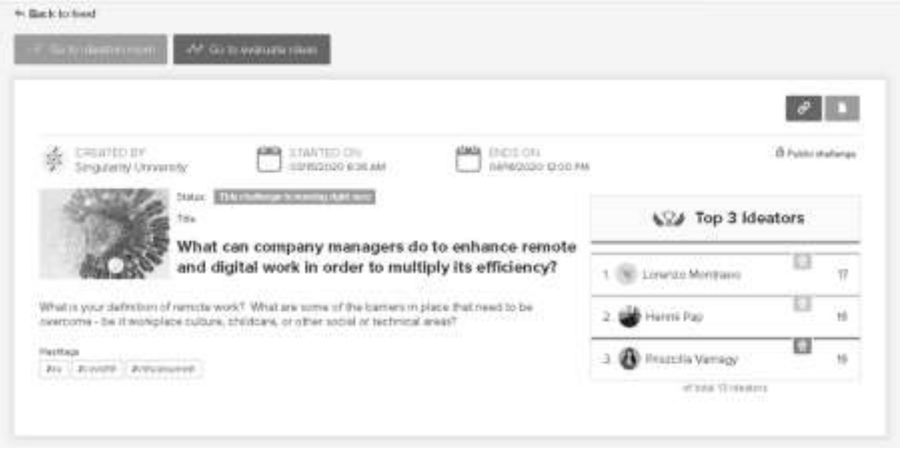
This challenge goes beyond geographical borders! See where ideators contributed from.

Figure 9

The geographical location of participants (www.be-novative.com)

By exporting the results to an Excel spreadsheet, a list is created that contains all the parameters and participants of the process. Sorting by the order of importance and thematic grouping of the proposals provides an opportunity to review, further weight, and then elaborate the implementation project in detail.

The ‘challenge’ interface is closed at the decision of the initiator upon completion or at a later date, and the summarized results can be viewed on the Be-novative interface. After running the program, on returning the initial interface, Figure 10 below is shown.



← Back to feed

[illegible] Go to evaluate ideas

CREATED BY
Singularity University

STARTED ON
[illegible]

ENDS ON
[illegible]

Public challenge

Status: [illegible]

Title

What can company managers do to enhance remote and digital work in order to multiply its efficiency?

What is your definition of remote work? What are some of the barriers in place that need to be overcome - be it workplace culture, childcare, or other social or technical areas?

Hashtags
[illegible]

Top 3 Ideators

1.	Lorenzo Montano	17
2.	Hanni Pap	16
3.	Piroska Varnagy	15

of total 13 ideators

Figure 10
Closed ‘challenge’ (www.be-novative.com)

Based on the above logic, our program was launched with the original research questions in mind.

4 Sample and Method

The research was initiated by a domestic KM organization. The opportunity was announced on the community’s website and newsletter, then the participants were randomly selected, based on registration. 25 people participated in the research (academic community representatives, entrepreneurs, project members, HR experts, business consultants, trainers). The research used two web interfaces in parallel, the Zoom room and the Be-novative program. At the announced time, participants entered the Zoom room. The head of the organization moderated the program here (goal formulation, program manual description, further information). The research process and criteria are presented in Table 1 below.

Table 1
The research logic

Research steps	Characteristics	Other
Announcement of the program	expert community website, newsletter	2 weeks before the program, KMEXPERT
Registration of participants	expert community website	KMEXPERT
Participant selection (sample)	randomly	25 persons
Participants' qualifications	representatives of the academic community, entrepreneurs, project members, HR professionals, business consultants, trainers	
The methodology used	Zoom room, be-novative program	In parallel
Moderator	Professional leader	Formulation of objectives, description of the use of the program, further information, conducting
Launching and running the program	Access to the Zoom room, registration on the be-novative platform	In parallel

The research question 'challenge' formulated to launch the program, which participants entered on the Be-novative interface, was: 'What knowledge management problems have emerged due to the situation caused by the spread of the coronavirus, what difficulties need to be coped with in organizations?' In the Zoom room, after the kick-off meeting, 3 small groups continued the joined-up thinking, according to the Be-novative logic. (Participants could choose to participate in a small group according to their interests, depending on which sub-topic they wanted to think about, to deepen the topic.) The sub-topics were:

What forms of direct (online) knowledge sharing should be developed that can be maintained even after the virus crisis? (The employees should receive answers to theirs as soon as possible, their sudden ideas should not be lost, and the lessons learned during work should be incorporated into the organization's knowledge assets.)

What online platform can be well applied to which knowledge management function? (Gathering online platforms that can help to solve the present situation and make knowledge management process steps more efficient in the long run.)

How to improve the on-boarding process of new entrants so that its elements can be used even after the crisis? (New employees entering the organization may start integration from home. How to adapt to the situation, what tools and methods are applicable that are useful in the long run?)

5 Results

The 3 sub-topics were developed in separate Zoom rooms, according to the participants' choice. (They could join the development of the topics based on their interest, see the number of evaluators in Table 2). The number of new ideas for each theme is also shown in Table 2. The above sample example shows that the program uses the Impact-Feasibility Graph (Figure 8; 11) to evaluate the votes of the participants (sorted by their importance), the results of which are also summarized in Table 2.

Following the 3 sub-topics, based on the above logic, 141 ideas were born. For the sub-topics, the program clarified the most important ideas (along with the scores given by the participants (Figure 7)), to which additional ones could be added to develop final solutions (Figure 6). Due to length constraints, only the first 3 ideas are listed in each sub-topic (Table 2).

Table 2
Sub-topic ideas and their evaluation scores (own construction)

Subtopic question	Ideas	Number of Evaluators	Number of new ideas	Evaluation scores
1.	Immediate access to important information from outside the corporate network	11	56	10
	Publication of professional guides	11	56	9.67
	Recording the date of regular meetings	11	56	9.56
2.	Can be used to store clouds, data	5	32	10
	Wunderlist, To-dos – task sharing	5	32	9.35
	Google Drive: editable table, opinions of many participants, easy to gather	5	32	9.35
3.	Development of an interactive E-learning interface that measures the value of progress and provides opportunities for feedback and questions	9	53	8.19
	Creating instructional videos	9	53	8
	Publishing company presentation materials	9	53	8

Based on the evaluation scores and the matrices generated by the program (effect vs. feasibility), the participants considered the first sub-topic worthy of further elaboration. The evaluation matrix generated by the program is shown in the figure below (Figure 11).

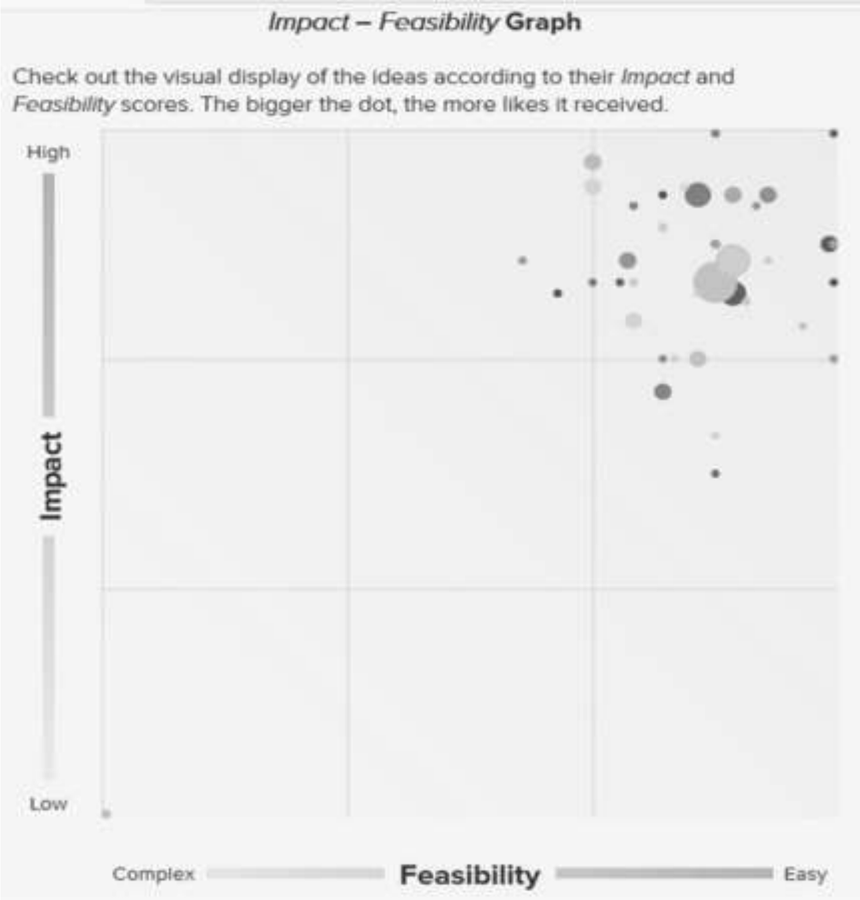


Figure 11

Evaluation matrix of the first sub-topic (www.be-novative.com)

Three of the 56 ideas in the first subtopic served as the basis for further joint work. (If necessary, more or fewer ideas can be elaborated on the decision of the organization's leadership. In this case, the number and content of ideas worth further elaboration were determined according to the joint decision of the professional circle.) The result provided by the program, which contained the grouped ideas, helped with the choice. Not only the overall score was decisive, but also the ideas rated by the participants as the most popular and most influential.

The possibility of combination and further elaboration on a common basis was also considered. Thus, of the 3 topics proposed for elaboration by the professional community (of the first 3 ideas, based on the evaluation), two were merged (entitled Knowledge transfer (online) and recording), the next topic renamed but keeping its original content (Scheduling online knowledge transfer), and an additional topic, considered important was included (Informal virtual meetings). During the discussion and further elaboration of the topics, several of the previous ideas were also included. Participants placed more emphasis on personal experiences, and, based on these, thought together about solutions for direct (online) knowledge exchange. In the form of voluntary commitment, in 2-3 minute phases, the participants presented the importance and the possibility of implementation of the ideas they proposed to be included. The topics processed in this way developed in a direction of more concrete, factual tasks during the joint discussion and conciliation. They progressed in a tunnel-like way in processing the topic, from the organizing principles to the implementation. In the following, the final proposed actions are briefly presented, giving an answer to the research question.

6 Discussion

Although there is a regulation in most organizations that prescribes the methodology of knowledge recording, in many cases, it is not up-to-date or clearly usable. Thus, accessing the necessary knowledge causes many problems. It is now possible (also in home-office) for employees to elaborate and keep up-to-date documents in their field, which can be accessed both internally and externally. The latter option is especially important for organizations that use closed internal systems, which until now could only be accessed from the organization's internal network [21-23].

2.3. Knowledge Transfer (Online) and Recording

Developing available and accessible knowledge repositories and ensuring their day-to-day use are only a part of the knowledge transfer practice. In the course of personal contacts, many quick questions, telephones, and emails in the workplace provide the ad hoc necessary information, which is currently difficult in the form of a home office. The main problems are the sharing of tacit knowledge and the lack of informal knowledge sharing that goes on as a part of unnoticed everyday practice. In the course of personal contacts, these knowledge-sharing solutions are embedded in the processes and they become a natural part of everyday work, which is not or very difficult to implement in the home-office format [24-26].

Given that the need for home-office (or at least, mixed/combined) working will remain after the pandemic situation has subsided, the forms and methods of knowledge transfer in the everyday practice of organizations need to be rethought and adapted to the changed needs.

2.4. Scheduling Online Knowledge Transfer

In order for employees to have daily, operative knowledge, daily, weekly meetings are essential. The timing and duration of these should be chosen so that further work is least disrupted. Although this idea seems to be evident, it turned out to be a serious problem in the first period of the arrangement of home office work (and its form is still undeveloped in many places) [27] [28]. Managers or staff would contact each other at all sorts of times, either with a problem, a question, or a report. These random, unscheduled online calls, emails, consultations interfere with focused thinking, continuous work, and concentration. This has been addressed and resolved in several places recently but, as the opinion of the expert community has shown, there are still many problematic cases of unplanned time management. The power of informal knowledge transfer, in addition to the formal, to sustain the community has been raised in several debates [28]. Therefore, the possibility of informal meetings was brought to focus as sub-topic 3.

2.5. Informal Virtual Meetings

Informal meetings, organized either in a regulated or an ad hoc manner, provide further possibilities for knowledge and information sharing. Just as personal coffee or beer together, informal meetings online serve the same aim. These virtual meetings, even in a "ceremonial" form, represent a serious cohesive force, culture formation, and knowledge transfer. Table 2 below summarizes the most important and easiest-to-apply options of the three sub-topics processed.

Table 2

Possibilities for solving subtopics (own construction)

Sub-topics	**Possible solutions**	**Implementation needs**	**Future sustaina bility**
1.	Process description, brief instructions, Q&A, briefings for new employees, responses to non-routine critical situations, answers to non-professional questions, online education, virtual coaching, mentoring, internal wiki, do's and don'ts, videos	Access to knowledge repositories through external communication, video, audio recording and transcription, recording of what is heard by the receiving party, managerial decision, and technical conditions	x

Sub-topics	Possible solutions	Implementation needs	Future sustaina bility
2.	"Stand up" morning start, starter 5-minute briefing, "One point lesson" in 10 minutes, "chunk of knowledge multiplier", gamification, "Virtual postcard method" modeled on "Share desk", "Teams", "Slack"	Setting the starting time and duration, facilitator, Provision of technical conditions and preparedness (education)	x
3.	virtual beer, random couple coffee, group chat (video chat, messenger, skype)	Attitude, openness, trust, provision of technical conditions	x Upgrade-able

Conclusions?

The last two years have changed the ways of thinking of and practicing knowledge management, and have raised its importance. Reviewing the previous months, valuable knowledge resources could be identified. The increase in digital content types, Zoom/Teams, and other forms of meetings, presentations and chat channel discussions has led knowledge management practitioners to find new ways of capturing and sharing this content so that it is easily accessible when needed. However, care must be taken not to overplay the potential of digitalization. Organizations must be able to maintain a reasonable level in terms of both the use of tools and the information they accumulate. It should be mentioned that the use and exploitation of organizational memory have also become important. In this context, the importance of organizational culture has also increased, which is crucial for the acceptance of knowledge management within the organization and for its structural and strategic positioning [23].

In recent months, it has become clear that a system of organizations with adequate knowledge existed well before the pandemic. These organizations were able to adapt relatively more quickly to the new digital working style compared to those that did not have a well-developed knowledge management system or the elements of it did not function (knowledge acquisition, sharing, storage, etc.). The pandemic has also accelerated the recognition in society of the significance of knowledge-sharing systems and the styles and importance of teleworking. In this context, the importance of trust in the virtual space has been enhanced. Several organizations have implemented coordinated techniques, which, in the future, will mean greater transparency in knowledge sharing. The future of organizations will continue to be cost-sensitive. According to Harold Koontz [29] bridging all areas of knowledge, linking relationships inside and outside the organization will be significant in the near future. Knowledge management and learning will take place online and virtually. This will require establishing an efficient knowledge center with artificial intelligence capabilities. After all, the pandemic has highlighted the importance of knowledge management worldwide.

The solution options listed in Table 3, in response to the raised problems, provide an opportunity to develop sustainable solutions in the future. The expert community continued to work in detail on the implementation options for the ideas raised, which are available on the professional organization's website [30]. Why have these solutions not been used so far? Often crisis situations bring out special problems and simple, cheap, and reasonable solutions. The balanced operation of the past has not forced professionals to recognize and deal with these solutions. PwC's research of 2019 [31] proved that leaders with crisis experience consider similar measures to be essential to underpin future success to the lessons listed at the beginning of this study. With the help of the Be-novative program, any organization can try and experience the power of joined-up thinking and crowdsourcing, the benefits of playfulness, time-saving solutions, and the possibility to overcome distances. The aim of this study is not to present scientific novelties that have not been announced so far but to present the application of a novel method, which (although it can be applied in any field of organizational life) tried to identify the difficulties of knowledge transfer and sharing in the field of knowledge management, which is a cardinal problem in every organizational operation, and to provide a contribution to their solution. Although the study does not go into the theoretical issues of knowledge management in any greater depth, it does offer a solution to the difficulties of its application by extending the possible methodology. Once the pandemic situation has subsided, the requirements of employees and new types of working conditions will certainly force organizations to rethink and reorganize their operating mechanisms, including the process and tools of knowledge management. This is supported by the method presented in the study and the results of the research.

The research topic is limited by the relatively small number of experts and the rapidly changing environmental conditions. In the future, it would be worth using the method described herein, to address similar questions and to compare the results longitudinally.

Acknowledgment

Thanks are due to all participants of the KMEXPERT professional community and the support of the Be-Novative Hungary Ltd. The publication was supported by the Pallas Athéné Domus Meriti Foundation.

References

[1] U. Hahn, D. Lagnado, S. Lewandowsky, and N. Chater, "Crisis knowledge management: Reconfiguring the behavioural science community for rapid responding in the Covid-19 crisis". PsyArXiv Preprint https://psyarxiv.com/hsxdk/ 2020, https://doi.org10.31234/osf.io/hsxdk

[2] M. Polanyi, *Personal Knowledge*. Chicago, USA: University of Chicago Press, 1958

[3] T. H. Davenport, and L. Prusak, *Working knowledge. How organizations manage what they know*. Massachusetts, USA: Harvard Business School Press, 1995

[4] I. Nonaka, and H. Takeuchi, (1995) *The knowledge creating company. How Japanese companies create the dynamics of innovation*. Oxford, UK: Oxford University Press, 1995

[5] G. D'Auria, and A. D. Smet, "Leadership in a crisis: Responding to the coronavirus outbreak and future challenges". McKinsey & Company. https://www.mckinsey.com/business-functions/organization/our-insights/leadership-in-a-crisis-responding-to-the-coronavirus-outbreak-and-future-challenges, (accessed Feb. 1, 2019)

[6] H. Ford, *My Life and Work: An Autobiography of Henry Ford*. Scotts Valley, California, USA: Createspace Independent Publishing Platform, 2015

[7] C. M. Giannantonio, and A. E. Hurley-Hanson *Extreme Leadership: Leaders, Teams and Situations Outside the Norm*. Cheltenham, UK: Edward Elgar, 2014

[8] Ch. Shakti and S. Tushar, „Knowledge Management Initiatives for Tackling the COVID-19 Pandemic in India", Metamorphosis, Vol. 20, No. 1, pp. 25-34, 2021, https://doi.org/10.1177/09726225211023677

[9] Briefing Knowledge Leaders: Paths to Greater Knowledge, Briefing April 2021, iManage [Online] Available: https://imanage.com/making-knowledge-work/?utm_source=briefingmag&utm_medium=offline&utm_campaign=mkw&vertical=large-law&utm_term=EMEA

[10] S. Ammirato, R. Linzalone and A. M. Felicetti,"Knowledge management in pandemics. A critical literature review", Knowl. Manage. Resour. & Prac. Vol. 19, No. 4, pp. 415-426, 2021, https://doi.org/10.1080/ 14778238.2020.1801364

[11] S. Surabhi, T. Nobin, and N. Ranjeet, "Knowledge sharing in times of a pandemic: An intergenerational learning approach", *Knowl. Proc. Manage. Spec. Iss.* Vol. 28, pp. 153-164, 2021, https://doi.org/10.1002/kpm.1669

[12] S. Ellis, Business Schools, Knowledge Management and COVID-19 – an early perspective. *Chartered Association of Business School*: https://charteredabs.org/business-schools-knowledge-management-and-covid-19-an-early-perspective/ (accessed Apr. 12, 2020)

[13] C. Scott,. The State of Knowledge Management in 2020. Apr. 27, Knowledge and Findability, [Online] Available: https://www.reworked.co/knowledge-findability/the-state-of-knowledge-management-in-2020/ (accessed Apr. 12, 2020)

[14] C. Luo, Y. Lan, X. Luo, H. Li, "The effect of commitment on knowledge sharing: an empirical study of virtual communities", *Technol. Forecast. Soc. Change*, Vol. 163, 2021, Art. no. 120438, https://doi.org/10.1016/j.techfore.2020.120438

[15] C. Speier, and J. Palmer, A definition of Virtualness. *Proc. Americas Conf. Information Systems*, 1998, pp. 571-573

[16] J. E. Klobas, *Becoming Virtual: Knowledge Management and Transformation of the Distributed Organization (Contributions to Management Science)* Heidelberg, Deutschland: Physica-Verlag, 2007

[17] W. Li, "Virtual knowledge sharing in a cross-cultural context". *J. Knowl. Manage*, Vol. 14, No. 1, pp. 38-50, 2010, https://doi.org/10.1108/13673271011015552

[18] R. D. Evans, J. X. Gao, N. Martin, and C. Simmonds, "A new paradigm for virtual knowledge sharing in product development based on emergent social software platforms". *J. Engin. Manufac*. Vol. 232, No. 13, pp. 2297-2308, 2018, https://doi.org/10.1177/0954405417699018

[19] T. Øystein, D. Amandeep and F. Bjørn-Tore, (2021) "Digital knowledge sharing and creative performance: Work from home during the COVID-19 pandemic", *Techn. Forecast. & Soc. Change* 170) Art. no. 120866 2021, https://doi.org/10.1016/j.techfore.2021.120866

[20] 13.08.2021 [Online] Available: www.be-novative.com

[21] A. S.-H. Lee, S. Wang, W. Yeoh, N. Ikasari, "Understanding the use of knowledge sharing tools", *J. Comput. Inf. Syst.* pp. 1-13, 2020, https://doi.org/10.1080/08874417.2020.1752850

[22] A. Rese, C. S. Kopplin, C. Nielebock, "Factors influencing members' knowledge sharing and creative performance in coworking spaces", *J. Knowl. Manage*. 2020, https://doi.org/10.1108/JKM-04-2020-0243

[23] A. Y. Lai, "Organizational collaborative capacity in fighting pandemic crises: A literature review from the public management perspective", *Asia-Pacific J. Publ. Health*, Vol. 24, No. 1, pp. 7-20, 2012, https://doi.org/10.1177/1010539511429592

[24] M. Généreux, M. Lafontaine, and A. Eykelbosh, "From Science to Policy and Practice: A Critical Assessment of Knowledge Management before, during, and after Environmental Public Health Disasters". *Int. J. Environ. Res. Publ. Health*, Vol. 16, No. 4, pp. 1-17, 2019, https://doi.org/10.3390/ijerph16040587

[25] L. Waizenegger, B. McKenna, W. Cai, T. Bendz, "An affordance perspective of team collaboration and enforced working from home during COVID-19". *Eur. J. Inf. Syst.* pp. 1-14, 2020, https://doi.org/10.1080/0960085X.2020.1800417

[26] M. Dorasamy, M. Raman, and M. Kaliannan, “Knowledge management systems in support of disasters management: A two decade review”. *Technol. Forecast. & Soc. Change,* Vol. 80, No. 9, pp. 1834-1853, 2013, https://doi.org/10.1016/j.techfore.2012.12.008

[27] J. Sua, Y. Yangb, and R. Duanc, “A CA-based heterogeneous model for knowledge dissemination inside knowledge-based organizations”. *J. Intell. & Fuzzy Syst.* Vol. 34, No. 4, pp. 2087-2097, 2018, https://doi.org/10.3233/JIFS-162116

[28] N. Van der Meulen, P. van Baalen, E. van Heck, and S. Mülder “No teleworker is an island: the impact of temporal and spatial separation along with media use on knowledge sharing networks”, *J. Inf. Technol.*, Vol. 34, No. 3, pp. 243-262, 2019, https://doi.org/10.1177/0268396218816531

[29] H. Koontz, C. O'Donnell, and H. Weihrich, *Essentials of management*, New York, USA: McGraw-Hill, 1986; New York, USA: Tata McGraw-Hill Education, 2010

[30] 12.12. 2020 [Online] Available: https://kmexper.hu

[31] PwC’s Global Crisis Survey [Online] Available: www.pwc.com/globalcrisissurvey (accessed Apr. 14, 2021)

Control Engineering Methods for Blood Glucose Levels Regulation

Jelena Tašić[1], Márta Takács[2] and Levente Kovács[1]

[1]Physiological Controls Research Center, Óbuda University, 1034 Budapest, Bécsi út 96/b, Hungary

[2]John von Neumann Faculty of Informatics, Óbuda University, 1034 Budapest, Bécsi út 96/b, Hungary

Email: tasic.jelena@uni-obuda.hu, takacs.marta@nik.uni-obuda.hu, kovacs@uni-obuda.hu

Abstract: In this article, we review recently proposed, advanced methods, for the control of blood glucose levels, in patients with type 1 diabetes. The proposed methods are based on various techniques, such as predictive control, filters, and machine learning. Results have shown that the artificial pancreas may control blood glucose levels better than conservative insulin administration, while avoiding the risk of hypoglycemia or hyperglycemia. The most commonly used methods for controlling blood glucose levels are giving good results, while methods based on machine learning algorithms also offer promising performance. Nevertheless, there are numerous challenges in designing algorithms for the artificial pancreas, which need to be considered. The aim of this research is to provide an overview of the latest achievements in this research field, find the best solutions and, ultimately, improve them in the future.

Keywords: Artificial pancreas; continuous glucose monitoring; model predictive control; sliding mode control; Kalman filters; machine learning; neural networks; type 1 diabetes

1 Introduction

In the last few decades, the number of people suffering from diabetes has constantly increased. According to the latest information, about 422 million people worldwide have diabetes, while 1.5 million deaths are directly attributed to diabetes each year [1]. Diabetes is a chronic autoimmune disease that occurs when the pancreas produces little or no insulin, as in type 1 diabetes (T1D), or when the body produces insulin but cannot use it effectively, as in type 2 diabetes (T2D). This disease destroys pancreatic β-cells, which are responsible for the production of the insulin peptide hormone, which regulates blood glucose (BG) levels.

Although T2D makes up 90-95% of cases, our focus will be on T1D, which is also known as juvenile diabetes or insulin-dependent diabetes.

Due to the lack of internal insulin production, patients with T1D need treatment with exogenous insulin, which is necessary for their survival. However, it should be taken into account that external administration of insulin also has its risks. In case of hypoglycemia, the patient’s BG level is below 3.9 mmol/l (70 mg/dl), which can lead to a potential loss of consciousness, seizures, coma, or death. On the other hand, we have elevated BG levels or hyperglycemia, where the patient’s BG level is greater than 11.1 mmol/l (200 mg/dl). This can lead to serious damage to the patient's body system and long-term complications such as neuropathies, nephropathy, or cardiovascular disease [2].

The artificial pancreas (AP) is a closed-loop glucose controller that provides automatic delivery of insulin. It consists of a continuous subcutaneous insulin infusion (CSII) pump which communicates with the continuous glucose monitoring (CGM) system that measures the BG levels, to automatically deliver insulin when needed [3]. After calculating the required amount of insulin, the pump releases and delivers an appropriate dose to the patient’s body using a specific control algorithm.

The results showed that the AP may control the BG levels and reduce the risk of hypoglycemia better than the conservative insulin administration compared with conventional insulin therapy (open-loop control) [4]. Even though the recently proposed methods give good results, there are many challenges in designing algorithms that need to be considered. Glucose metabolic disorders can occur under the influence of various factors such as changes in diet, circadian rhythm, stress, alcohol consumption, unannounced physical exercise, menstrual cycle, chronic metabolic variations, or insulin sensitivity [5]. Also, there are additional factors such as urgent time requirements, unknown analytical relationships between custom parameters and measured values, and security issues, which present additional challenges for the development of the algorithms [6].

After a brief introduction of T1D and AP, in Section II we present a review of control methods based on model predictive control, Bayesian optimization, sliding mode control, proportional integral derivative control, linear parameter varying, iterative learning control, active disturbance rejection control, robust fixed point transformations, disturbance observer, terminal synergetic, state feedback linearization, and bioinspired AP. In Section III is given a brief review of a method for the identification of parameters, while in Section IV we review methods based on kernel and Kalman filters. In Section V we present a review of novel approaches based on machine learning such as unsupervised and supervised learning, clustering, artificial neural networks, and bioinspired reinforcement learning. Finally, we conclude with Section VI.

2 Approaches Based on Control Methods

In this Section, we present recently proposed methods for controlling and regulating BG levels, but also for preventing large delays in insulin absorption. One of the most commonly used methods is model predictive control, which is used to handle meal announcements [7], control BG levels, limit insulin infusion rates, and improve its delivery through the prediction horizon. The sliding mode control approach is a common method used for handling insulin stabilization, regulating glycaemia, and improving glucose regulation. A proportional integral derivative model-based approach that relies on physiological models that consider the operation of metabolism is also commonly used. Other approaches include Bayesian optimization, linear parameter varying, iterative learning control, robust fixed point transformations, active disturbance control, disturbance observer, terminal synergetic, state feedback linearization, and bioinspired AP.

2.1 Model Predictive Control Approach

Most of the proposed methods, which have been tested in clinical studies, are based on the linear model predictive control (MPC). MPC has shown that it is able to stabilize BG levels, but also to improve the bolus calculator for more efficient meal management [8-10]. Currently, used calculators depend on the correction between BG levels and insulin intakes. The reason is that a linear relationship between the size of the announced meal and the insulin bolus should be assumed [11].

While Chakrabarty et al. [12] used an observer-based MPC algorithm with the novel event-triggered communication (ETC) method for reducing sensor-controller transmissions, Cairoli et al. [3] improved MPC with a signal temporal logic (STL) method using the Hovorka compartment ordinary differential equation (ODE) model (Fig. 1). The STL was able to provide safe BG pathways allowing soft constraints, even during meal disturbances, while avoiding hypoglycemia and hyperglycemia.

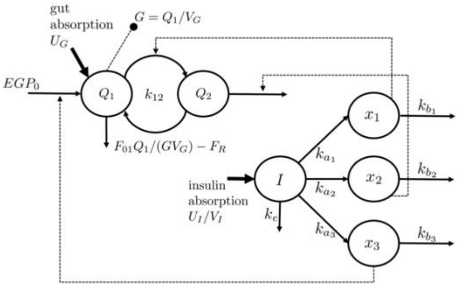


Figure 1

Scheme of the applied Hovorka compartment ODE model [3]

A recursive subspace-based empirical modeling algorithm based on the predictor-based subspace identification (PBSID) method was presented by Rashid et al. [13] to determine the linear dynamic model, while CGM measurements were used to determine the appropriate values for the plasma insulin concentration (PIC) bounds and risk indexes. The proposed method provided a stable, time-varying, and individualized state-space model for predicting CGM measurements, while keeping BG levels within the safe range, without meal announcement.

Boiroux et al. [14] presented the identified physiological model for describing the glucose-insulin dynamics for the nonlinear MPC (NMPC), where virtual patients were generated using the Hovorka model, as well as its parameter distributions, to test the identification procedure (Fig. 2). The results showed that the proposed method has the potential to be used in NMPC algorithms.

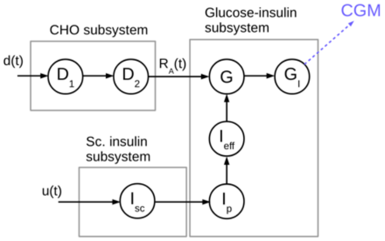


Figure 2
The proposed MPV model [14]

On the other hand, Embaby et al. [15] proposed a novel adaptive NMPC (AMPC) approach, consisting of a Cobelli model, a fuzzy logic controller (FLC), a feedforward neural network (FFNN), and an adaptation method, for BG levels regulation. The FLC was used to compute the amount of insulin infusion and maintain BG levels in a normal range, while the genetic algorithm was used to solve FLC optimization problems and improve search performance. The FFNN was used as the NMPC to manage the insulin delay between the time of injection and its interaction, while the adaptation method was used to adjust the compensation of the proposed system for physiological differences between patients. The results indicated that the time of increase in BG levels was in the normal range, causing less hyperglycemia.

To update the real-time control penalty parameters for a zone MPC (ZMPC) method, Shi et al. [16] applied a dynamic cost function. The proposed method gave a good performance for announced moderate meal-bolus, unannounced meals, and physical exercises, and improved BG levels, while the rate of insulin delivery was within a safe range, without the risks of hypoglycemia.

Chakrabarty et al. [17] implemented an embedded ZMPC method, using the fast adaptive memetic algorithm (FAMA) and the fast alternating direction method of multipliers (FADMM) algorithm to solve convex constraints of the linear MPC method. The generated closed-loop data were used to select the optimization algorithm and the appropriate setting parameters. The proposed method was able to maintain BG regulation and it was compatible with other embedded systems.

Abuin et al. [18] improved the robustness of a time-varying pulsatile ZMPC (pZMPC) with the linear time-invariant (LTI) method, by adding a circadian insulin sensitivity (S_I) scheme. The performance of the time-varying pZMPC was compared with respect to the linear time-invariant pZMPC-LTI, with the models configured with low and high S_I. The pZMPC-h achieved better performance during high S_I intervals by improving the analyzed metrics, while during the period of low S_I it produced hyperglycemic events.

Hajizadeh et al. [19] integrated a multivariable AP (mAP) method with a controller performance monitoring, assessment, and modification (CPMAM) system to analyze closed-loop behavior, modify MPC parameters, and automate insulin delivery systems during different meal amounts and exercise times (Fig. 3). The CPMAM system was proposed for the adaptive learning MPC (AL-MPC) and then applied in the mAP system for real-time estimation using various key performance indexes (KPIs). The control of BG levels was improved without the risk of hypoglycemia.

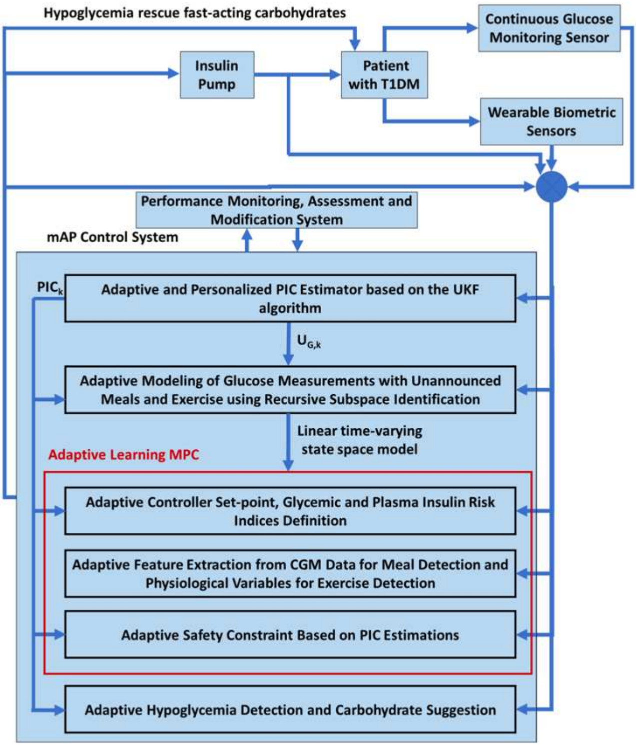

Hypoglycemia rescue fast-acting carbohydrates
Continuous Glucose Monitoring Sensor
Insulin Pump
Patient with T1DM
Wearable Biometric Sensors
Performance Monitoring, Assessment and Modification System
mAP Control System
PIC$_k$
Adaptive and Personalized PIC Estimator based on the UKF algorithm
U$_{G,k}$
Adaptive Modeling of Glucose Measurements with Unannounced Meals and Exercise using Recursive Subspace Identification
Linear time-varying state space model
Adaptive Learning MPC
Adaptive Controller Set-point, Glycemic and Plasma Insulin Risk Indices Definition
Adaptive Feature Extraction from CGM Data for Meal Detection and Physiological Variables for Exercise Detection
Adaptive Safety Constraint Based on PIC Estimations
Adaptive Hypoglycemia Detection and Carbohydrate Suggestion

Figure 3

The proposed mAP method with integrated CPMAM system [19]

Reenberg et al. [20] presented a linear MPC-based algorithm for critically ill patients in an intensive care unit (ICU). The proposed algorithm is based on a stochastic continuous-discrete state-space model and represents a model of multi-input single-output (MISO) transfer function. To demonstrate the performance of the closed-loop algorithm, the Bergman minimal model (BMM), the Hovorka ICU Model, and the Chase ICU Model were used. Additional measurement delays, which are associated with glucose-sensing or enteral nutrition, have made it difficult to achieve strict glycemic control, which increases the risk of hypoglycemia.

Sun et al. [21] proposed a novel event-triggered MPC (ET-MPC) algorithm for personal insulin dosing to regulate BG levels and reduce computational requirements during unannounced meals and physical activity, performed according to pre-established criteria. The proposed method proved to be robust to a CGM data deficiency and signal loss, providing personalized assessment, while maintaining BG levels in a safe range without risk of hypoglycemia.

2.2 Bayesian Optimization Approach

A method based on the multivariate Bayesian optimization (BO) approach and the dynamic parameter selection module for solving the parameter adaptation problem was presented by Shi et al. [6]. The dynamic parameter selection module was used to determine the parameters, while the BO-based optimization module was used to automatically adjust the selected parameter and to optimize an unknown cost function, as is shown in Fig. 4. The efficiency and robustness of the proposed algorithm was verified in two scenarios. In the first case, the rate of insulin delivery was improved, while BG levels were reduced to the euglycemic range. In the second case, the algorithm was able to improve the duration of insulin delivery. Therefore, the proposed method may properly adjust the parameters to achieve their regulation, without the risk of hypoglycemia.

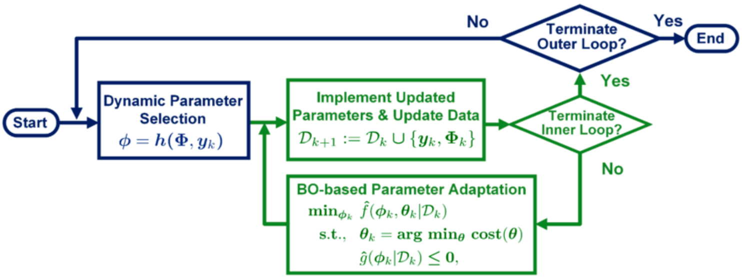

No
Terminate Outer Loop?
Yes
End
Yes
Start
Dynamic Parameter Selection
$\phi = h(\Phi, y_k)$
Implement Updated Parameters & Update Data
$\mathcal{D}_{k+1} := \mathcal{D}_k \cup \{y_k, \Phi_k\}$
Terminate Inner Loop?
No
BO-based Parameter Adaptation
$\min_{\phi_k} \hat{f}(\phi_k, \theta_k | \mathcal{D}_k)$
s.t., $\theta_k = \arg\min_\theta \text{cost}(\theta)$
$\hat{g}(\phi_k | \mathcal{D}_k) \leq 0,$

Figure 4

The proposed method based on the dynamic parameter selection module (blue) and the optimization module (green) [6]

2.3 Sliding Mode Control Approach

Beneyto et al. [22] applied an insulin-only controller using fast-acting carbohydrates (CHO) for the recommender system to improve the regulation of BG levels caused by unannounced physical activity. The proposed method consists of a proportional–derivative (PD) controller with insulin feedback (IFB) and a safety auxiliary feedback element (SAFE) layer, as shown in Fig. 5. The SAFE layer consists of insulin on board (IOB) constraints, a sliding mode reference conditioning (SMRC) block, and a low-pass first-order filter, while the CHO controller is based on a predictive quantified PD controller. Comparison of the original insulin-only controller and the combined insulin CHO recommender system showed that the novel combined system may reduce daily episodes of hypoglycemia and increase the rate of insulin delivery within acceptable limits.

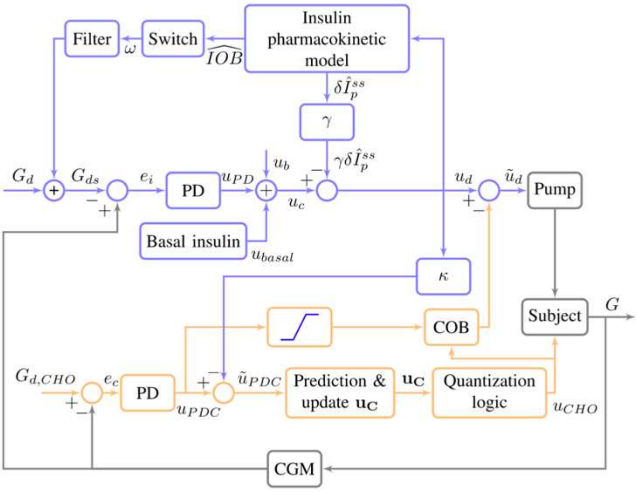


Figure 5

The proposed insulin-only controller (blue) with the CHO controller (orange) [22]

Moscardö et al. [23] used the SMRC method to improve the coordinated configuration (CC) control structure with IOB limitation for coordinated BG control levels (Fig. 6). A comparison of CC and CC-SMRC control structures was made based on meals, snacks, and exercise scenarios. Although the results of the proposed method were better during the exercise periods, than during the meals, in the most demanding exercise scenario, insulin delivery levels were not sufficient to prevent hypoglycemia.

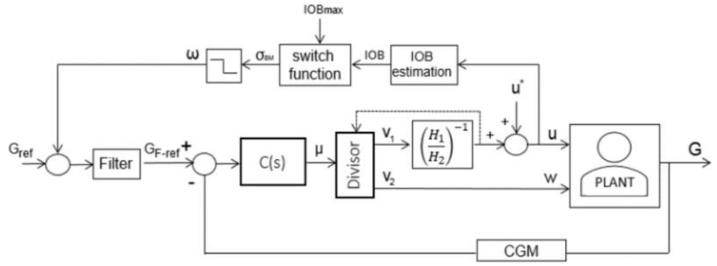


Figure 6
The proposed method based on the CC-SMRC controller [23]

Leyva et al. [24] presented methods based on the positive sliding mode control (SMC) and the control Lyapunov function (CLF), where the cascade structure of the physiological model was used to improve the rate and stabilize BG levels, while the compartmental mathematical model was used to reproduce glucose metabolism, and insulin and glucagon dynamics. Although both methods managed to solve the problem of stabilization, the CLF gave better results by improving the convergence rate and generating a continuous signal that prevented the accumulation of insulin.

A finite-time synergistic control approach based on a gain-scheduled Luenberger observer (GSLO) was presented by Alam et al. [25] to establish a closed-loop insulin delivery system (Fig. 7). A finite-time back-stepping SMC strategy was used to regulate glycemia, while the CLF law was systematically achieved in a recursive procedure. The intravenous glucose tolerance test (IVGTT) model (BMM), was considered to design a nonlinear control algorithm. The robustness of the system was achieved despite external disturbances, while postprandial hyperglycemia and hypoglycemia were suspended.

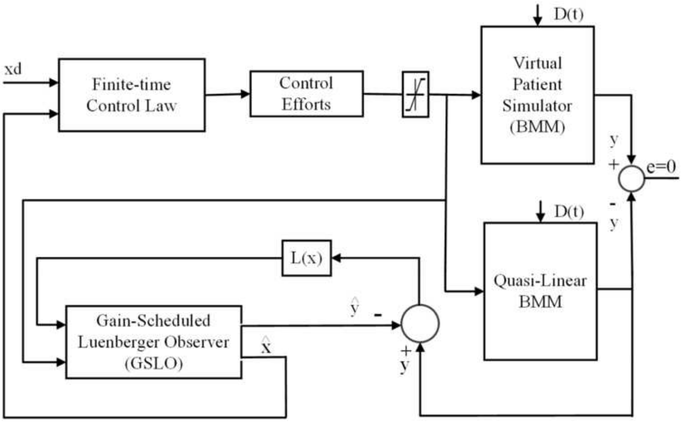


Figure 7
The proposed closed-loop control system based on the GSLO [25]

2.4 Proportional Integral Derivative Control Approach

Kushner et al. [26] presented a novel non-deterministic data-driven model with a proportional integral derivative (PID) based closed-loop system to predict patient reaction to the proposed system while maintaining BG levels control. The proposed model was able to efficiently adjust key controller parameters and improve BG levels control. To reduce insulin absorption delay, Barnes and Jones [27] applied the continuous intraperitoneal insulin infusion (CIPII) method based on the PID controller. The IMC-PID controller based on the internal model control (IMC) tuning method was introduced, which employs an inverter to realize the PID controller feedback. The time delay was adjusted using a first-order with time delay (FOPTD) model, along with a Pade approximation. The proposed controller was able to successfully control the oscillations of BG levels.

A novel PID control-based method, consisting of an adaptive weighted PID (AWPID) controller and a look-ahead PID with retrospective estimation error correction (LAPID-REC), was presented by Alshalalfah et al. [28] to prevent large delays incurred in insulin action and glucose sensitivity. In the AWPID approach, the proportional gain of the PID controller was rated based on the short-term CGM history, while in the LAPID-REC approach prospective estimates of future measurements were used to calculate the control action with retrospective estimation error correction. The safety and performance of standard PID control were improved, while the LAPID-REC approach showed high performance over existing techniques, especially under sensor noise, counteracting the long delays that occur in CGM and insulin action.

2.5 Linear Parameter Systems Approach

Eigner et al. [29] presented an advanced controller design method for a physiological model, using a theorem based on the linear parameter varying (LPV) and linear matrix inequality (LMI), which was applied on a modified version of the minimal model. The resulting controller used a state feedback type control rule due to the applicable LPV-LMI conditions.

Conversely, Colmegna et al. [30] extended the IOB safety loop method with an inner switched LPV (SLPV) controller and an outer sliding-mode safety layer (SAFE), to limit the controller’s action, during multiple meals and exercises. A mode selection algorithm was added to combine the hyperglycemia detection module with heart rate (HR) data for automatically adjusted controller settings (Fig. 8). The proposed method was able to effectively reduce the risk of hypoglycemia during the moderate exercise scenario.

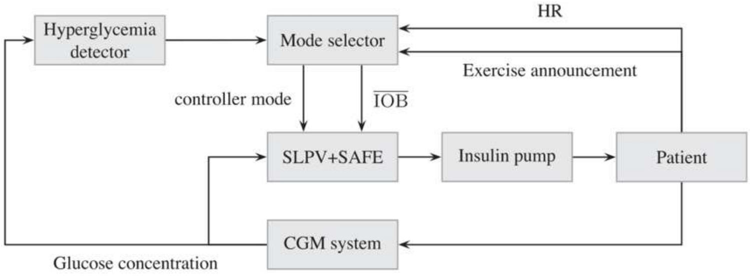


Figure 8

The extended method based on the SLPV and SAFE [30]

2.6 Iterative Learning Control Approach

Modifications of the Dalla Man metabolic model were proposed by Cescon et al. [31] by adding a long-acting insulin absorption model to facilitate validation of the control strategy for the multiple daily injections (MDI) therapy. A once-a-day iterative learning control (ILC) based dosing method was proposed to provide basal insulin delivery. Fig. 9 presents the proposed model of subcutaneous insulin absorption, with the amount of injected rapid-acting and long-acting. In the case of fasting, meal and meal with induced insulin resistance, the ILC performed better than the open-loop dose, by providing an appropriate amount of basal insulin.

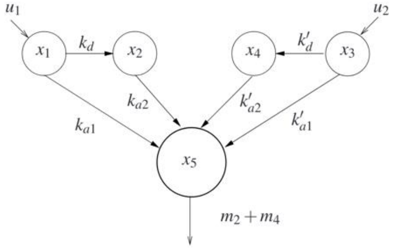


Figure 9

The proposed model of subcutaneous insulin absorption [31]

Cescon et al. [32] also proposed the ILC algorithm for the delivery of long-acting (basal) and rapid-acting (bolus) insulin, for patients following the MDI therapy (Fig. 10). The ILC updates basal therapy consisting of one long-acting insulin injection per day, while by updating the mealtime-specific insulin-to-carbohydrate ratio, the run-to-run (R2R) controller adjust meal bolus therapy. The results showed that the proposed method can provide robustness against random variations, resistance to protocol deviations while improving glycemic regulation over time.

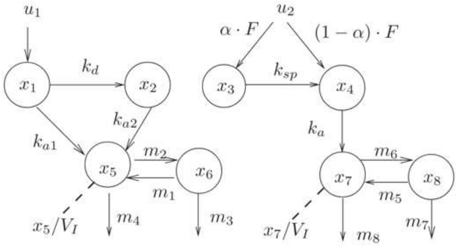


Figure 10

The proposed compartment model of insulin subsystem [32]

2.7 Robust Fixed Point Transformations Approach

To create a robust and adaptive control approach for BG levels control, Kovács et al. [33] presented a novel robust fixed point transformation (RFPT) based controller approach which consists of the two delay blocks corresponding to the cycle time of the digital controller (Fig. 11). Although the proposed method constantly absorbed external glucose concentration, it was able to interfere with the negative effect of inherent model uncertainties and measurement disturbances, while reducing the risk of hyperglycemia and hypoglycemia.

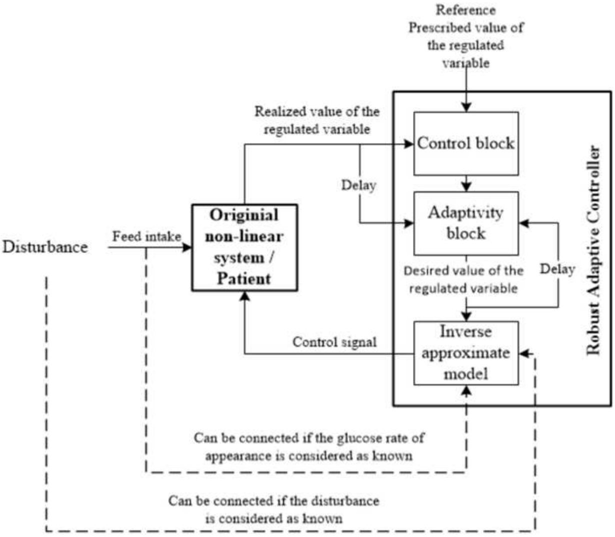

Reference
Prescribed value of the regulated variable
Realized value of the regulated variable
Control block
Delay
Adaptivity block
Robust Adaptive Controller
Disturbance
Feed intake
Originial non-linear system / Patient
Desired value of the regulated variable
Delay
Inverse approximate model
Control signal
Can be connected if the glucose rate of appearance is considered as known
Can be connected if the disturbance is considered as known

Figure 11

Scheme of the proposed RFPT method [33]

2.8 Active Disturbance Rejection Control Approach

Cai et al. [34] proposed an active disturbance rejection control (ADRC) method by adding IOB and insulin delivery constraints to ensure the safety of the control algorithm. The controller consists of the ADRC module (composed of tracking differentiator, extended state observer (ESO) and nonlinear feedback) and the constraints module (composed of the IOB, non-negative and maximum input constraints). The proposed method was able to achieve satisfactory performance of BG regulation and insulin delivery rate without the risk of hypoglycemia.

2.9 Disturbance Observer Approach

Sanz et al. [35] used disturbance observer (DOB) to estimate the effect of unannounced meals, and feedforward compensator for the insulin pharmacokinetics, to control postprandial BG levels of patients. The results showed that the DOB may successfully estimate and counteract the effect of meals and the sudden drops in BG levels while avoiding hypoglycemia. For unannounced meals with high CHO content, a median time-in-range was 80% with large intra-subject variability, while for announced meals the median time-in-range was increased up to 88%, even considering severe bolus mismatch and CHO counting errors.

2.10 Terminal Synergetic and Feedback Linearization Controller Approaches

Babar et al. [36] extended BMM (EBMM) with the nonlinear terminal synergetic controller (TSC) and the state feedback linearization based controller (SFC), while the Lyapunov theory was used to provide asymptotic stability of the proposed controllers. White noise was added to the EBMM, and then the performance of each controller was evaluated to check their ability to withstand disturbance. Compared to other controllers, the TSC gave the best results with about zero steady-state error, lesser settling, convergence time, with acceptable overdrafts.

2.11 Bioinspired AP Approach

A bi-hormonal bioinspired AP (BiAP) controller was extended with a novel hybrid hormonal-insulin sensitivity glucose (InSiG) by Güemes et al. [37], to determine insulin and glucagon doses with the coordinated bi-hormonal BiAP controller, and to determine the desired S_I from CGM with a standard PD (sPD) controller. After comparing the InSiG controller and the coordinated bi-hormonal BiAP controller, the results showed that the InSiG controller was able to improve BG levels control while maintaining within the target range without the risk of hypoglycemia.

Although, the proposed controller was able to reduce the delivered dose of insulin and significantly reduce the glucagon dose, the relationship between the magnitude of nervous system stimulation and the S_I dynamics remained unknown.

3 Approaches Based on Sensitivity Analysis

Staal et al. [38] investigated methods to improve recognition and estimation of the most appropriate model parameters to reduce the parameters of critical models. The identification of nonlinear state-space model parameters was also investigated. The nonlinear observability rank condition (NORC) was used for structural, while sensitivity analysis and the Fisher information matrix (FIM) were used for practical identifications. A simplified model, derived from CGM, scarce self-monitoring of BG (SMBG), meal and insulin data, showed to be useful for the AP applications.

4 Approaches Based on Filters

In this Section, we review recently proposed methods based on extended Kalman and kernel filtering algorithms for detecting unannounced meals or missed meal announcements, real-time insulin pump faults detection, insulin infusion rate regulation, and BG levels control.

4.1 Extended Kernel Filter

To improve computational efficiency in online glucose prediction, Yu et al. [39] extended an adaptive kernel filter (KRLS) algorithm with the sparsification criteria. The KRLS algorithm was combined with the approximate linear dependency (ALD) and the surprise criterion (SC) to design an online sparse ALD-KRLS and SC-KRLS algorithms. The proposed online adaptive method proved to be insensitive to abnormal or inaccurate CGM measurements and it was adaptable to prediction models. Thus, it could effectively reduce the computational load and regulate the time delay in the nonlinear dynamics of glucose.

4.2 Extended Kalman Filter

Fushimi et al. [40] proposed the integration of the automatic switching signal generator (SSG) into the automatic regulation of glucose (ARG) algorithm and an advanced version of the switched linear quadratic Gaussian (SLQG) controller, to regulate the basal insulin infusion rate. The SSG module, based on the KF, was used to generate a filtered version of BG levels. Despite the large delay in selecting the post-meal controller mode, the proposed algorithm had efficiency of 83.3% in terms of meal detection, it was able to regulate the basal insulin infusion rate and generate insulin feedback during unannounced meals, without significantly increasing the risk of hypoglycemia or hyperglycemia.

A novel kernel function for the Gaussian process was proposed by Ortmann et al. [2] by improving the existing MPC controller and solving the problem of noise in measurements during unannounced meals. The unscented KF was used to assess the condition, extract data, and change S_I. The extracted data were processed using a Gaussian filter to predict future effects, while the MPC optimized the received data to calculate the volume of insulin injections, as shown in Fig. 12. The collected training data became insensitive to noise after the application of the Gaussian process, making the controller insensitive to unannounced meals.

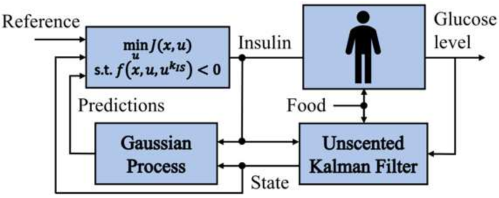


Figure 12

The proposed method based on the unscented KF, Gaussian process, and MPC [2]

To present a novel adaptive model-based algorithm for detecting unannounced meals, Fathi et al. [41] used a linear KF to compute the evaluation of BG measurements, applying the statistical generalized likelihood ratio test under the null hypothesis, to estimate the impact of an unannounced meal on BG levels. The proposed algorithm managed to successfully detect unannounced moderate meals 96.29% of the time, without false positives.

Boiroux et al. [42] presented a model for nonlinear estimation of the maximum probability of estimated parameters, where the state covariance matrix and its gradient were calculated using explicit Runge-Kutta schemes, while the method implementation was verified by using a numerical example for nonlinear parameter estimation.

On the other hand, Kovács et al. [43] applied advanced LPV, linear matrix inequality (LMI), tensor product (TP) model transformation, and extended KF (EKF) control methods, to guarantee strong safety control of BG levels. An extension of the minimal model was applied to simulate the glucose-insulin dynamics and glucose and insulin absorption. The control structure of the TP model was combined with LMI based optimization and LPV control (TP-LMI-LPV controller), EKF, and D/A converter (Fig. 13). The proposed controller was able to intervene effectively during the process and provide appropriate control actions, thus satisfy predefined requirements while avoiding hypoglycemia.

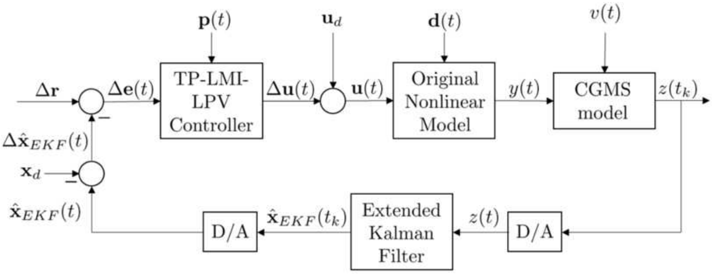


Figure 13

Scheme of the proposed method with the TP-LMI-LPV controller and mixed EKF [43]

Kovács et al. [44] also introduced the dual EKF (DEKF) framework to estimate the state variables and model parameters at the same time by utilizing the discrete LPV methodology. A nonlinear model was applied to the quasi-LPV (qLPV) model (derived from the nonlinear Cambridge T1DM model) to map the noise effects that occurred during the application of the CGM system. The results showed that the proposed method was able to estimate state variables with good accuracy.

Meneghetti et al. [45] proposed a method for real-time insulin pump fault detection and missed meal announcements to improve the safety of the AP system architecture. The proposed method consists of an offline model and a predictor module, and an online prediction and alert module, as shown in Fig. 14. The confounding factor introduced by meals was tested to detect insulin pump faults ability. The proposed method was able to improve patient feedback, providing various alarms and effectively preventing pump malfunctioning due to user errors, without causing hyperglycemic events.

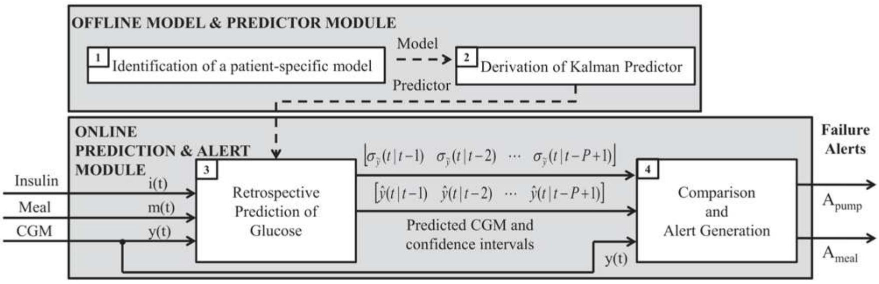

OFFLINE MODEL & PREDICTOR MODULE
1 Identification of a patient-specific model
Model
Predictor
2 Derivation of Kalman Predictor
ONLINE PREDICTION & ALERT MODULE
Insulin
Meal
CGM
i(t)
m(t)
y(t)
3 Retrospective Prediction of Glucose
$[\sigma_{\hat{y}}(t|t-1)\ \sigma_{\hat{y}}(t|t-2)\ \cdots\ \sigma_{\hat{y}}(t|t-P+1)]$
$[\hat{y}(t|t-1)\ \hat{y}(t|t-2)\ \cdots\ \hat{y}(t|t-P+1)]$
Predicted CGM and confidence intervals
y(t)
4 Comparison and Alert Generation
Failure Alerts
A_{pump}
A_{meal}

Figure 14

The proposed fault detection method [45]

Sala-Mira et al. [46] compared the LPV dual KF, the LPV joint KF, and the nonlinear sliding mode observer (NSMO), to evaluate the effect of observer structure on estimation performance. Observers were composed of the Hovorka and Identifiable Virtual Patient (IVP) models, which represents a compromise between the Bergman and Hovorka model in terms of structural complexity and accuracy. Analysis of variance (ANOVA) and multiple comparisons were used to assess the individual factors. Based on PIC and rate of appearance, the results showed that proportions of variance were low for each factor, indicating a small difference between observer structures.

5 Machine Learning Algorithms

In this Section, we review recently proposed Machine Learning methods [47] for automatic insulin infusion, insulin pump failure detection, physical activity prediction, overnight glycemic control quality prediction, online prediction of BG levels and its stability, gradient problems, but also to improve prediction accuracy and robustness of previous methods. The proposed methods are based on unsupervised and supervised learning, clustering, artificial neural networks, and bioinspired reinforcement learning.

5.1 Algorithms Based on Unsupervised Learning

An unsupervised model-free approach based on data-driven techniques for anomaly detection was presented by Meneghetti et al. [48] to detect insulin pump malfunction. Machine learning (ML) methods for detecting anomalies, using local outlier factor (LOF), connectivity-based outlier factor (COF), and isolation forest (iF/iForest), were applied to the extracted set of features. To overcome correlations between time-closed samples, the for time series data (4TSD) procedure was applied to LOF and COF. The optimal parameter configuration for LOF and iForest was able to provide satisfactory detection performance while maintaining high accuracy. After comparison with the traditional multivariate control chart (MCC) method, the results showed that COF outperformed other methods, while LOF and iForest offered comparable performance. Despite the good performance, iForest has been shown to be prone to errors and instabilities.

5.2 Algorithms Based on Supervised Learning

Güemes et al. [49] proposed a novel data-driven method for predicting the overnight quality of glycemic control, by analyzing a small data set from CGM measurements, meal intake, and insulin bolus. To classify the overnight quality of glycemic control, binary classification algorithms such as random forest classifier (RFC), artificial neural networks (ANN), support vector machine (SVM), linear logistic regression (LLR), and extended tree classifier (ETC) were used. The proposed method was able to predict overnight BG levels within the target range with reasonable accuracy of 0.7. However, a larger data set is needed to fully validate the proposed method.

The solution based on supervised ML, to predict future BG levels, was proposed by Eigner et al. [50]. To prove the concept, TensorFlow and Keras frameworks were used with the AIDA diabetes simulator for data generation. The results showed that the proposed method gives an accurate prediction of BG levels within acceptable limits, with overall accuracy of 0.879, taking into account that the accuracy of predicting normal BG levels should be improved.

Dénes-Fazakas et al. [51] applied synthetic data generated by an extended open-source version of the Jacobs T1DM simulator, which employs the Cambridge model and contains an embedded physical activity sub-model. To predict the presence of physical activity, a logistic regression, AdaBoost classifier, decision tree classifier, Gaussian naive Bayes, the k-nearest neighbor classifier (k-NN), SVM, RFC, and multilayer perceptron networks (MLP) were used, and then trained classifiers were applied to all feature vectors of the test data set. Decision tree, k-NN, and RFC gave the best results, with overall accuracy of 0.91, 0.95 and 0.98. Other models may be also suitable, but they need additional mechanisms to avoid false positives.

5.3 Algorithms Based on Clustering

Montaser et al. [52] proposed a seasonal autoregressive integrated moving average (SARIMAX) model, an extended version of the non-seasonal ARIMAX model, and examined the possibility of preprocessing original CGM measurements to obtain sets of similar glycemic profiles (clusters) to identify a seasonal model of postprandial periods. Using the fuzzy c-means (FCM) clustering method, the number of sets and corresponding features of the BG profile was obtained in the modeling step, while the Box-Jenkins methodology was used to identify the seasonal model for each cluster set. The results showed that using online BG predictions through a global seasonal model may reduce the risk of hypoglycemia or hyperglycemia.

A data-driven approach for determining the final set of daily CGM profiles (motifs) was presented by Lobo et al. [53] so that almost every generated daily profile could be matched with one of the motifs from the final set. A training data set was used to identify candidate motif sets, while a validation data set was used to select the final set. The results showed that robustness was successfully established while matching with representative daily CGM profiles in the test data set was 99.0%.

5.4 Algorithms Based on Artificial Neural Networks

Aliberti et al. [54] applied a nonlinear autoregressive (NAR) neural network and long short-term memory (LSTM) on BG signals, to improve prediction accuracy and robustness of previous methods (Fig. 15). NAR was used to solve BG stability problems, while LSTM was used to explode and disappear the gradient, as well as to maintain long-term information over time.

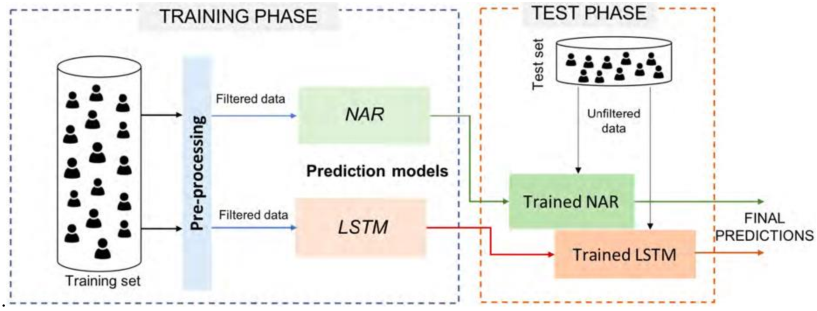


Figure 15
The proposed solution with applied NAR and LSTM methods [54]

Compared to the recurrent neural networks (RNNs), the LSTM was more resistant to the exploding and vanishing gradient problems. The NAR model gave good prediction accuracy only for a short-term period (30-minute prediction horizon), while the LSTM exhibited very good performance for predicting both short-term and long-term BG levels (60 minute prediction horizon).

Li et al. [55] proposed a convolutional RNN (CRNN) method that consists of a multilayer convolutional neural network (CNN), a RNN layer with LSTM cells, and fully connected layers, to predict BG levels. The CNN was used to extract features or patterns of the multidimensional time series, while a modified RNN was used to analyze the previous sequential data and predict BG levels. The results showed that the proposed method was able to predict BG levels with high accuracy.

To predict BG levels, Zhu et al. [56] proposed a novel deep learning framework with the edge inference on a microcontroller unit (MCU) embedded in a low-power system, by using CGM measurements and the RNN that builds on LSTM (Fig. 16). Collected data from wearable devices were uploaded to the server. Then, a well-trained deep neural network (DNN) was embedded in the MCU and further implemented in wearable devices to help in decision making. The proposed framework was agnostic to the types of neural networks employed and learning targets, and it showed a good BG prediction performance. Therefore, it could be applied for the realization of various tasks on wearable devices, such as event detection (e.g. meals, exercise, illness, errors) and glucose regulation.

A novel deep reinforcement learning (RL) model for optimizing single-hormone (insulin) and dual-hormone (insulin and glucagon) delivery was presented by Zhu et al. [57].

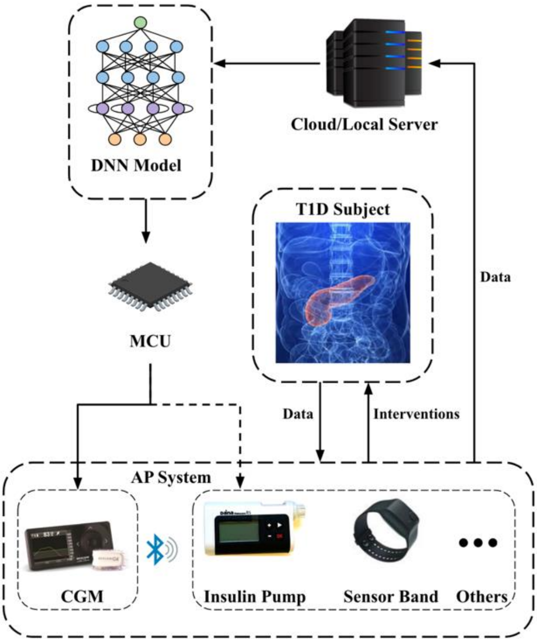


Figure 16
The system architecture of the proposed DNN-based method [56]

Dilated RNNs were applied to the structure of double deep Q-network (DQNs), to develop personalized models through a two-step framework that involves transfer learning (Fig. 17). Proposed methods gave good control of BG levels with a significant reduction in hypoglycemia making the use of deep RL a sustainable approach to closed-loop BG control, with the best TIR score of 93%.

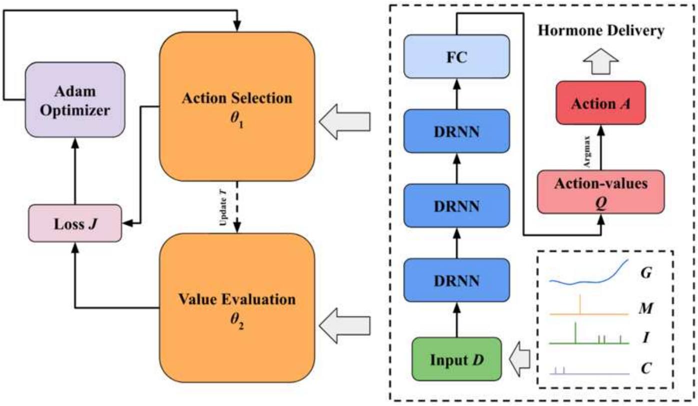


Figure 17
Scheme of the proposed double DQN method [57]

5.5 Bioinspired Reinforcement Learning

A novel AI-based bioinspired RL approach for automated insulin infusion was proposed by Lee et al. [58], to maintain BG levels and robustness of the CGM sensor. The layer-wise relevance propagation (LRP) method was used to analyze input-output relevance and define the rate of insulin infusion. The proposed LRP method was able to provide information about insulin distribution, making the decision step by step, without distinguishing between basal and bolus insulin, which is similar to the principle of human β-cells. A trained policy could automatically maintain fasting BG levels after unannounced meal intake without a prediction model, automatically respond and regulate postprandial glucose, provide robustness with respect to CGM sensor noise, achieve a mean BG level in the normal range of 89.56%, and without the risks of hypoglycemia.

Conclusions

In this work, we reviewed various recently proposed methods, based on predictive control, sensitivity analysis, filters and machine learning algorithms, intended for regulating insulin delivery and controlling BG levels in patients with T1D. The control approaches included control methods based on model predictive control, Bayesian optimization, sliding mode control, proportional integral derivative control, linear parameter varying, iterative learning control, active disturbance rejection control, robust fixed point transformation, disturbance observer, terminal synergetic controller, state feedback linearization based controller and bi-hormonal bioinspired AP. Combining common control methods has shown good results in controlling BG levels while maintaining a safe range. The proposed methods based on the Kalman filter, combined with different control methods, gave good results in state variables and model parameters estimation.

Other successful approaches included methods that are based on machine learning techniques, such as, unsupervised and supervised learning, clustering, artificial neural networks and bioinspired reinforcement learning. The Long Short-term Memory has shown very good performance for predicting short-term and long-term BG levels, while combining with recurrent neural networks could predict BG levels with high accuracy. Novel, deep reinforcement machine learning algorithms, promise improved performance for larger experimental datasets, with the support of powerful hardware platforms. Clustering methods gave good results in predictive modeling, decision support, and automated systems, while bioinspired reinforcement learning was able to provide insulin distribution information, automated postprandial regulation, sensor robustness, and fully automate BG levels control for unannounced meals. For the case of insulin infusion, bioinspired reinforcement learning made the decisions step by step, without distinguishing between basal and bolus insulin, similar to the principle of the human β-cells.

Acknowledgment

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program (grant agreement No. 679681). Project No. 2019-1.3.1-KK-2019-00007 has been implemented with the support provided from the National Research, Development, and Innovation Fund of Hungary, financed under the 2019-1.3.1-KK funding scheme.

References

[1] World Health Organization. (2020) Diabetes fact sheet. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/diabetes

[2] L. Ortmann, D. Shi, E. Dassau, F. J. Doyle, B. J. E. Misgeld, and S. Leonhardt, "Automated insulin delivery for type 1 diabetes mellitus patients using Gaussian process-based model predictive control," in *Proc. 2019 Amer. Control Conf. (ACC)*, Philadelphia, PA, USA, Jul. 2019, pp. 4118-4123.

[3] F. Cairoli, G. Fenu, F. A. Pellegrino, and E. Salvato, "Model predictive control of glucose concentration based on signal temporal logic specifications," in *Proc. 2019 6th Int. Conf. Control, Decision Inform. Technol. (CoDIT)*, Paris, France, Apr. 2019, pp. 714-719

[4] J. Tašić, G. Eigner, and L. Kovács, "Review of algorithms for improving control of blood glucose levels," in *Proc. 2020 IEEE 18th Int. Symp. Intell. Syst. Inform. (SISY)*, Subotica, Serbia, Sept. 2020, pp. 179-184

[5] R. A. DeFronzo, E. Ferrannini, P. Zimmet, and K. G. M. M. Alberti, *International Textbook of Diabetes Mellitus*, 4th ed. Oxford, UK: Wiley-Blackwell, 2015

[6] D. Shi, E. Dassau, and F. J. Doyle, "A multivariate Bayesian optimization framework for long-term controller adaptation in artificial pancreas," in *Proc. 2018 IEEE Conf. Decision Control (CDC)*, Miami Beach, FL, USA, Dec. 2018, pp. 276-283

[7] P. Szcześniak, G. Tadra, and Z. Fedyczak, "Model predictive control of hybrid transformer with matrix converter," *ACTA Polytechnica Hungarica*, Vol. 17, No. 1, pp. 25-40, Jan. 2020

[8] R. Hovorka, J. M. Allen, D. Elleri, L. J. Chassin, J. Harris, D. Xing, C. Kollman, T. Hovorka, A. M. F. Larsen, M. Nodale, A. D. Palma, M. E. Wilinska, C. L. Acerini, and D. B. Dunger, "Manual closed-loop insulin delivery in children and adolescents with type 1 diabetes: a phase 2 randomised crossover trial," *The Lancet*, Vol. 375, No. 9716, pp. 743-751, Feb. 2010

[9] S. Schmidt, D. Boiroux, A. K. Duun-Henriksen, L. Frøssing, O. Skyggebjerg, J. B. Jørgensen, N. K. Poulsen, H. Madsen, S. Madsbad, and

K. Nørgaard, "Model-based closed-loop glucose control in type 1 diabetes: the DiaCon experience," *J. Diabetes Sci. Technol.*, Vol. 7, No. 5, pp. 1255-1264, Sep. 2013

[10] S. D. Favero, D. Bruttomesso, F. D. Palma, G. Lanzola, R. Visentin, A. Filippi, R. Scotton, C. Toffanin, M. Messori, S. Scarpellini, P. Keith-Hynes, B. P. Kovatchev, J. H. DeVries, E. Renard, L. Magni, A. Avogaro, and C. Cobelli, "First use of model predictive control in outpatient wearable artificial pancreas," *Diabetes Care*, Vol. 37, No. 5, pp. 1212-1215, May 2014

[11] S. Schmidt and K. Nørgaard, "Bolus calculators," *J. Diabetes Sci. Technol.*, Vol. 8, No. 5, pp. 1035-1041, May 2014

[12] A. Chakrabarty, S. Zavitsanou, F. J. Doyle, and E. Dassau, "Model predictive control with event-triggered communication for an embedded artificial pancreas," in *Proc. 2017 IEEE Conf. Control Technol. Appl. (CCTA)*, Mauna Lani, HI, USA, Aug. 2017, pp. 536-541

[13] M. Rashid, I. Hajizadeh, and A. Cinar, "Predictive control with variable delays in plasma insulin action for artificial pancreas," in *Proc. 2018 IEEE Conf. Decision Control (CDC)*, Miami Beach, FL, USA, Dec. 2018, pp. 291-296

[14] D. Boiroux, Z. Mahmoudi, and J. B. Jørgensen, "Parameter estimation in type 1 diabetes models for model-based control applications," in *Proc. 2019 Amer. Control Conf. (ACC)*, Philadelphia, PA, USA, Jul. 2019, pp. 4112-4117

[15] A. A. Embaby, Z. Nossair, and H. Badr, "Adaptive nonlinear model predictive control algorithm for blood glucose regulation in type 1 diabetic patients," in *Proc. 2020 2nd Novel Intell. Leading Emerg. Sci. Conf. (NILES)*, Giza, Egypt, Oct. 2020, pp. 109-115

[16] D. Shi, E. Dassau, and F. J. Doyle, "Adaptive zone model predictive control of artificial pancreas based on glucose- and velocity-dependent control penalties," *IEEE Trans. Biomed. Eng.*, Vol. 66, No. 4, pp. 1045-1054, Apr. 2019

[17] A. Chakrabarty, E. Healey, D. Shi, S. Zavitsanou, F. J. Doyle, and E. Dassau, "Embedded model predictive control for a wearable artificial pancreas," *IEEE Trans. Control Syst. Technol.*, Vol. 28, No. 6, pp. 2600-2607, Nov. 2020

[18] P. Abuin, J. E. Sereno, A. Ferramosca, and A. H. Gonzalez, "Closed-loop MPC-based artificial pancreas: handling circadian variability of insulin sensitivity," in *Proc. 2020 Argentine Conf. Automat. Control (AADECA)*, Buenos Aires, Argentina, Oct. 2020, pp. 1-6

[19] I. Hajizadeh, N. Hobbs, M. Sevil, M. Rashid, M. R. Askari, R. Brandt, and A. Cinar, "Performance monitoring, assessment and modification of an

adaptive MPC: automated insulin delivery in diabetes," in *Proc. 2020 Eur. Control Conf. (ECC)*, St. Petersburg, Russia, May 2020, pp. 283-288

[20] A. T. Reenberg, D. Boiroux, T. K. Skovborg Ritschel, and J. Bagterp Jørgensen, "Model predictive control of the blood glucose concentration for critically ill patients in intensive care units," in *Proc. 2019 IEEE 58th Conf. Decision Control (CDC)*, Nice, France, Dec. 2019, pp. 3762-3769

[21] X. Sun, M. Rashid, M. R. Askari, N. Hobbs, R. Brandt, and A. Cinar, "Event-triggered decision support and automatic control systems for type 1 diabetes," in *Proc. 2021 IEEE EMBS Int. Conf. Biomed. Health Inform. (BHI)*, Athens, Greece, Jul. 2021, pp. 1-4

[22] A. Beneyto, A. Bertachi, J. Bondia, and J. Vehi, “A new blood glucose control scheme for unannounced exercise in type 1 diabetic subjects,” *IEEE Trans. Control Syst. Technol.*, Vol. 28, No. 2, pp. 593-600, Mar. 2020

[23] V. Moscardö, P. Herrero, J. L. Diez, M. Giménez, P. Rossetti, and J. Bondia, “In silico evaluation of a parallel control-based coordinated dual-hormone artificial pancreas with insulin on board limitation,” in *Proc. 2019 Amer. Control Conf. (ACC)*, Philadelphia, PA, USA, Jul. 2019, pp. 4759-4764

[24] H. Leyva, G. Quiroz, F. A. Carrillo, and R. Femat, “Insulin stabilisation in artificial pancreas: a positive control approach,” *IET Control Theory Appl.*, Vol. 13, No. 7, pp. 970-978, Apr. 2019

[25] W. Alam, Q. Khan, R. A. Riaz, R. Akmeliawati, I. Khan, and K. S. Nisar, "Gain-scheduled observer-based finite-time control algorithm for an automated closed-loop insulin delivery system," *IEEE Access*, Vol. 8, pp. 103088-103099, May 2020

[26] T. Kushner, D. Bortz, D. M. Maahs, and S. Sankaranarayanan, “A data-driven approach to artificial pancreas verification and synthesis” in *Proc. 2018 ACM/IEEE 9th Int. Conf. Cyber-Physical Syst. (ICCPS)*, Porto, Portugal, Apr. 2018, pp. 242-252

[27] A. J. Barnes and R. W. Jones, “PID-based glucose control using intra-peritoneal insulin infusion: an in silico study,” in *Proc. 2019 14th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Xi’an, China, Jun. 2019, pp. 1057-1062

[28] A. -L. Alshalalfah, G. B. Hamad, and O. A. Mohamed, "Towards safe and robust closed-loop artificial pancreas using improved PID-based control strategies," *IEEE Trans. Circuits Syst. I, Reg. Papers*, Vol. 68, No. 8, pp. 3147-3157, Aug. 2021

[29] G. Eigner, I. Böjthe, A. Mészáros, and L. Kovács, "Robust H∞ controller design for T1DM based on relaxed LMI conditions," in *Proc. 2019 IEEE 23rd Int. Conf. Intell. Eng. Syst. (INES)*, Gödöllő, Hungary, Apr. 2019, pp. 000363-000368

[30] P. H. Colmegna, F. D. Bianchi, and R. S. Sánchez-Peña, "Automatic glucose control during meals and exercise in type 1 diabetes: proof-of-concept in silico tests using a switched LPV approach," *IEEE Control Syst. Lett.*, Vol. 5, No. 5, pp. 1489-1494, Nov. 2021

[31] M. Cescon, S. Deshpande, F. J. Doyle, and E. Dassau, "Iterative learning control with sparse measurements for long-acting insulin injections in people with type 1 diabetes," in *Proc. 2019 Amer. Control Conf. (ACC)*, Philadelphia, PA, USA, Jul. 2019, pp. 4746-4751

[32] M. Cescon, S. Deshpande, R. Nimri, F. J. Doyle III, and E. Dassau, "Using iterative learning for insulin dosage optimization in multiple-daily-injections therapy for people with type 1 diabetes," *IEEE Trans. Biomed. Eng.*, Vol. 68, No. 2, pp. 482-491, Feb. 2021

[33] L. Kovács, G. Eigner, B. Czakó, M. Siket, and J. K. Tar, "An opportunity of using robust fixed point transformation-based controller design in case of type 1 diabetes mellitus," in *Proc. 2019 1st Int. Conf. Societal Autom. (SA)*, Krakow, Poland, Sept. 2019, pp. 1-7

[34] D. Cai, J. Song, J. Wang, and D. Shi, "Glucose regulation for subjects with type 1 diabetes using active disturbance rejection control," in *Proc. 2019 Chin. Control Conf. (CCC)*, Guangzhou, China, Jul. 2019, pp. 6970-6975

[35] R. Sanz, P. García, J. -L. Díez, and J. Bondia, "Artificial pancreas system with unannounced meals based on a disturbance observer and feedforward compensation," in *IEEE Trans. Control Syst. Technol.,* Vol. 29, No. 1, pp. 454-460, Jan. 2021

[36] S. A. Babar, I. A. Rana, I. S. Mughal, and S. A. Khan, "Terminal synergetic and state feedback linearization based controllers for artificial pancreas in type 1 diabetic patients," *IEEE Access*, Vol. 9, pp. 28012-28019, Feb. 2021

[37] A. Güemes, P. Herrero, and P. Georgiou, "A novel glucose controller using insulin sensitivity modulation for management of type 1 diabetes," in *Proc. 2019 IEEE Int. Symp. Circuits Syst. (ISCAS)*, Sapporo, Japan, May 2019, pp. 1-5

[38] O. M. Staal, A. L. Fougner, S. Sælid, and Ø. Stavdahl, "Glucose-insulin metabolism model reduction and parameter selection using sensitivity analysis," in *Proc. 2019 Amer. Control Conf. (ACC)*, Philadelphia, PA, USA, Jul. 2019, pp. 4104–4111

[39] X. Yu, M. Rashid, J. Feng, N. Hobbs, I. Hajizadeh, S. Samadi, M. Sevil, C. Lazaro, Z. Maloney, E. Littlejohn, L. Quinn, and A. Cinar, "Online glucose prediction using computationally efficient sparse Kernel filtering algorithms in type-1 diabetes," *IEEE Trans. Control Syst. Technol.*, Vol. 28, No. 1, pp. 3-15, Jan. 2020

[40] E. Fushimi, P. Colmegna, H. D. Battista, F. Garelli, and R. Sánchez-Peña, “Unannounced meal analysis of the ARG algorithm,” in *Proc. 2019 Amer. Control Conf. (ACC)*, Philadelphia, PA, USA, Jul. 2019, pp. 4740-4645

[41] A. E. Fathi, E. Palisaitis, B. Boulet, L. Legault, and A. Haidar, “An unannounced meal detection module for artificial pancreas control systems,” in *Proc. 2019 Amer. Control Conf. (ACC)*, Philadelphia, PA, USA, Jul. 2019, pp. 4130-4135

[42] D. Boiroux, T. K. S. Ritschel, N. Kjølstad Poulsen, H. Madsen, and J. B. Jørgensen, "Efficient computation of the continuous-discrete extended Kalman filter sensitivities applied to maximum likelihood estimation," in *Proc. 2019 IEEE 58th Conf. Decision Control (CDC)*, Nice, France, Dec. 2019, pp. 6983-6988

[43] L. Kovács, G. Eigner, M. Siket, and L. Barkai, "Control of diabetes mellitus by advanced robust control solution," *IEEE Access*, Vol. 7, pp. 125609-125622, Aug. 2019

[44] L. Kovács, M. Siket, I. Rudas, A. Szakál, and G. Eigner, "Discrete LPV based parameter estimation for TIDM patients by using dual extended Kalman filtering method," in *Proc. 2019 IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Bari, Italy, Oct. 2019, pp. 1390-1395

[45] L. Meneghetti, A. Facchinetti, and S. D. Favero, “Model-based detection and classification of insulin pump faults and missed meal announcements in artificial pancreas systems for type 1 diabetes therapy,” *IEEE Trans. Biomed. Eng.*, Vol. 68, No. 1, pp. 170-180, Jan. 2021

[46] I. Sala-Mira, M. Siket, L. Kovacs, G. Eigner, and J. Bondia, “Effect of model, observer and their interaction on state and disturbance estimation in artificial pancreas: an in-silico study,” *IEEE Access*, Vol. 9, pp. 143549-143563, Oct. 2021

[47] C. M. Bishop, *Pattern Recognition and Machine Learning*. Secaucus, NJ, USA: Springer-Verlag, 2006

[48] L. Meneghetti, M. Terzi, S. D. Favero, G. A. Susto, and C. Cobelli, “Data-driven anomaly recognition for unsupervised model-free fault detection in artificial pancreas,” *IEEE Trans. Control Syst. Technol.*, Vol. 28, No. 1, pp. 33-47, Jan. 2020

[49] A. Güemes, G. Cappon, B. Hernandez, M. Reddy, N. Oliver, P. Georgiou, and P. Herrero, “Predicting quality of overnight glycaemic control in type 1 diabetes using binary classifiers,” *IEEE J. Biomed. Health Inform.*, Vol. 24, No. 5, pp. 1439-1446, May 2020

[50] G. Eigner, M. Nagy, and L. Kovács, "Machine learning application development to predict blood glucose level based on real time patient data," in *Proc. 2020 RIVF Int. Conf. Comput. Commun. Technol. (RIVF)*, Ho Chi Minh City, Vietnam, Oct. 2020, pp. 1-6

[51] L. Dénes-Fazakas, L. Szilágyi, J. Tasic, L. Kovács, and G. Eigner, "Detection of physical activity using machine learning methods," *2020 IEEE 20th Int. Symp. Comput. Intell. Inform. (CINTI)*, Budapest, Hungary, Nov. 2020, pp. 167-172

[52] E. Montaser, J. -L. Díez, P. Rossetti, M. Rashid, A. Cinar, and J. Bondia, "Seasonal local models for glucose prediction in type 1 diabetes," *IEEE J. Biomed. Health Inform.*, Vol. 24, No. 7, pp. 2064-2072, Jul. 2020

[53] B. Lobo, L. Farhy, M. Shafiei, and B. Kovatchev, "A data-driven approach to classifying daily continuous glucose monitoring (CGM) time series," *IEEE Trans. Biomed. Eng*., Vol. 69, No. 2, pp. 654-665, Feb. 2022

[54] A. Aliberti, I. Pupillo, S. Terna, E. Macii, S. D. Cataldo, E. Patti, and A. Acquaviva, “A multi-patient data-driven approach to blood glucose prediction,” *IEEE Access*, Vol. 7, pp. 69311-69325, May 2019

[55] K. Li, J. Daniels, C. Liu, P. Herrero, and P. Georgiou, “Convolutional recurrent neural networks for glucose prediction,” *IEEE J. Biomed. Health Inform.*, Vol. 24, No. 2, pp. 603-613, Feb. 2020

[56] T. Zhu, L. Kuang, K. Li, J. Zeng, P. Herrero, and P. Georgiou, "Blood glucose prediction in type 1 diabetes using deep learning on the edge," in *Proc. 2021 IEEE Int. Symp. Circuits Syst. (ISCAS)*, Daegu, Korea, May 2021, pp. 1-5

[57] T. Zhu, K. Li, P. Herrero, and P. Georgiou, "Basal glucose control in type 1 diabetes using deep reinforcement learning: an in silico validation," *IEEE J. Biomed. Health Inform.*, Vol. 25, No. 4, pp. 1223-1232, Apr. 2021

[58] S. Lee, J. Kim, S. W. Park, S. M. Jin, and S. M. Park, “Toward a fully automated artificial pancreas system using a bioinspired reinforcement learning design: in silico validation,” *IEEE J. Biomed. Health Inform.*, Vol. 25, No. 2, pp. 536-546, Feb. 2021

Application of Artificial Immune Networks in Continuous Function Optimizations

Petar Čisar[1], Sanja Maravić Čisar[2], Brankica Popović[1], Kristijan Kuk[1], Igor Vuković[3]

[1]University of Criminal Investigation and Police Studies, Cara Dušana 196, 11080 Zemun, Serbia

[2]Subotica Tech – College of Applied Sciences, Marka Oreškovića 16, 24000 Subotica, Serbia

[3]Ministry of the Interior of the Republic of Serbia, Kneza Miloša 101, 11000 Belgrade, Serbia

petar.cisar@kpu.edu.rs, sanjam@vts.su.ac.rs, brankica.popovic@kpu.edu.rs, kristijan.kuk@kpu.edu.rs, igor.vukovic@mup.gov.rs

Abstract: This paper deals with the application of artificial immune networks in continuous function optimizations. The performance of the immunological algorithms is analyzed using the Optimization Algorithm Toolkit. It is shown that the CLIGA algorithm has, by far, the fastest convergence and the best score - in terms of the number of required iterations, for the analyzed continuous function. Also, based on the test results, it was concluded, that the lowest total number of iterations for the defined run time was achieved with the opt-IA algorithm, followed by the CLONALG and CLIGA algorithms.

Keywords: artificial immune networks; Optimization Algorithm Toolkit; continuous function optimization; performance

1 Introduction

Optimization is defined as the procedure for determining the best set of acceptable conditions, to achieve a specific objective and is formulated in mathematical terms. Optimization problems arise in a broad area of real-world applications. This paper presents the specificities of artificial immune-based algorithms to solve continuous optimization problems. The problem of continuous optimization includes determining the extremes of a function of one or many real variables, within a value spectrum, with possible constraints.

The main contribution made by this study is the introduction of methods for implementing the Optimization Algorithm Toolkit (OAT) environment, in order to examine the performance of artificial immune networks in continuous function optimizations.

This work is structured as follows: Section 2 in the form of related work offers an outlining the general principle of how this immune algorithm functions and explaining the categorization of artificial immune system (AIS) algorithms. Section 3 describes the problems of continuous function optimization and presents the implemented test functions. This is followed by Section 4, with the focus on performance measuring of artificial immune algorithms in a suitable software environment. Finally, in Section 5, conclusions are drawn and future study inquiries are suggested.

2 Related Work

The passages below give ample background information so as to enable the comprehension of the immunological algorithms.

There are several definitions for artificial immune systems. One of them was formulated by de Castro and Timmis [1]: artificial immune systems are adaptive systems, inspired by theoretical immunology and observed immune functions, principles and models.

In an artificial immune network (system) a set of integral components called B cells (B lymphocytes, binary-encoded candidate solutions), interact with each other and go through certain cloning and mutation operations. Similar to artificial neural networks, as shown by Dragulescu and Albu [22], artificial immune networks can learn new information and use previously learned information.

The appropriate theoretical approach, as well as different applications of AIS are elaborated in [14-16].

The general functioning principle of the immune algorithm is outlined in the flowchart [2].

An immune algorithm mathematically models the immune diversity, network theory and clonal selection as a multi-modal function optimization problem.

The authors de Castro and von Zuben [3], formulated the functional similarities and differences between the immune system and immune algorithm.

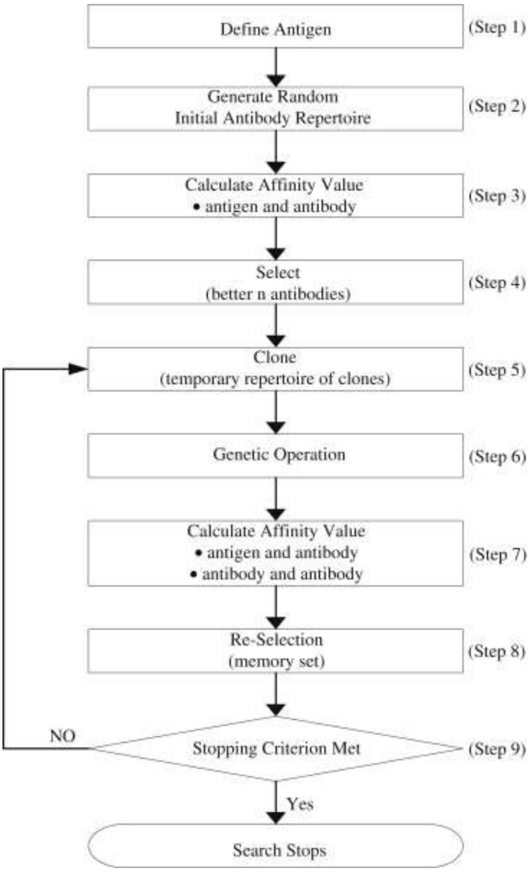


Figure 1
Flowchart of the immune algorithm [2]

Table 1
Immune system vs. immune algorithm

Immune system	**Immune algorithm**
Antigen	Problem to be solved
Antibody	Best solution vector
Recognition of antigen	Identification of the problem
Production of antibody from memory cells	Recalling a past successful solution
Lymphocyte differentiation	Maintenance of good solutions (memory)
T-cell suppression	Elimination of surplus candidate solutions
Proliferation of antibody	Use of genetic operators to create new antibodies

Lopez, Morales and Niño [4] concluded that four major AIS algorithms were under constant development: Negative Selection Algorithms (NSA), Artificial Immune Networks (AINE), Clonal Selection Algorithms (CLONALG) and Dendritic Cell Algorithms (DCA). AIS algorithms can be successfully applied in problems related to clustering, data visualization, control, pattern recognition (intrusion detection [23-27]) as well as, various types of optimizations.

Cutello and Nicosia pointed out in [5] that a simple clonal selection algorithm was named Immunological Algorithm (IA) and later was renamed to Simple Immune Algorithm (SIA). This algorithm analyzes a population of antibodies (B cells) that are exposed to a clonal expansion process. The process involves the cloning of cells with the implementation of a hypermutation parameter.

Table 2

SIA description

Parameter	**Description**
P	Population of antibodies
l	Length of binary string representation
d	Antibody population size
dup	Duplication of the bit string
clone	The number of clones created for each antibody
hypermutation	Modification of a bit string (bit flipping), requires the specification (ρ) of the probability of flipping each bit

The original Immunological Algorithm (IA) was renamed and represented many times by different authors. Other names included Simple Immune Algorithm (SIA), Cloning, Information Gain, Aging (CLIGA), and Optimization Immune Algorithm (opt-IA, opt-IMMALG).

The functioning of SIA can be elaborated by the following pseudocode, as formulated by Brownlee [6]:

```
P <- rand(d, l)
ForEach p of P Do               //presentation
      affinity(p)
EndFor
While Not StopCondition Do
      ForEach p of P Do         //clonal expansion
            Pc <- clone(p, dup)
      EndFor
      ForEach c of Pc Do        //affinity maturation
            hypermutate(c)
      EndFor
      ForEach c of Pc Do        //presentation
            affinity(c)
      EndFor
```

```
P <- select(Pc, P, d)        //clonal selection
EndWhile
```

Listing 1
Simple Immune Algorithm - pseudocode

An AIS combining CLONALG (one of the immune algorithms proposed for pattern recognition and optimization) with the immune network theory resulted in a model named aiNet [17]. This model was successfully applied to data compression and clustering applications, including non-linear separable and high-dimensional problems. The optimization version of aiNet (opt-aiNet) algorithm [12] [13] was applied to several uni/bi - dimensional functions in order to assess its performance. The results illustrated its behavior for some of the problems tested and compared it with results obtained by CLONALG. Three functions were used for testing: the multi-modal function, roots, and Schaffer's function. It was demonstrated that the opt-aiNet located 61 local maximums, while the CLONALG located 18. In addition, the opt-aiNet positions one single individual in each peak, which can overcome the 'waste of resources' disadvantage of the CLONALG.

Ulutas and Kulturel-Konak formulated the general steps of CLONALG [7]:

1) Initialization - randomly initialize a population
2) Evaluation - given a collection of patterns to identify, determine the match of each pattern with each member in the population
3) Selection and cloning - select a few of the best affinity elements and clone (duplicate) them according to their affinity with the antigen
4) Hypermutation - all the clones should be mutated proportional to the affinity with the input sample
5) Editing receptors - add the mutated individuals to the population and reselect a number of the maturated (optimized) individuals as memory
6) Steps 2-5 should be repeated until a stopping criterion is reached

As a representative of AIS algorithms, CLIGA algorithm uses three parameters:

Table 3
CLIGA description

Parameter	Description
d	Antibody population size
dup	Duplication of the bit string
τ_b	B cell's expected mean life (aging operator)

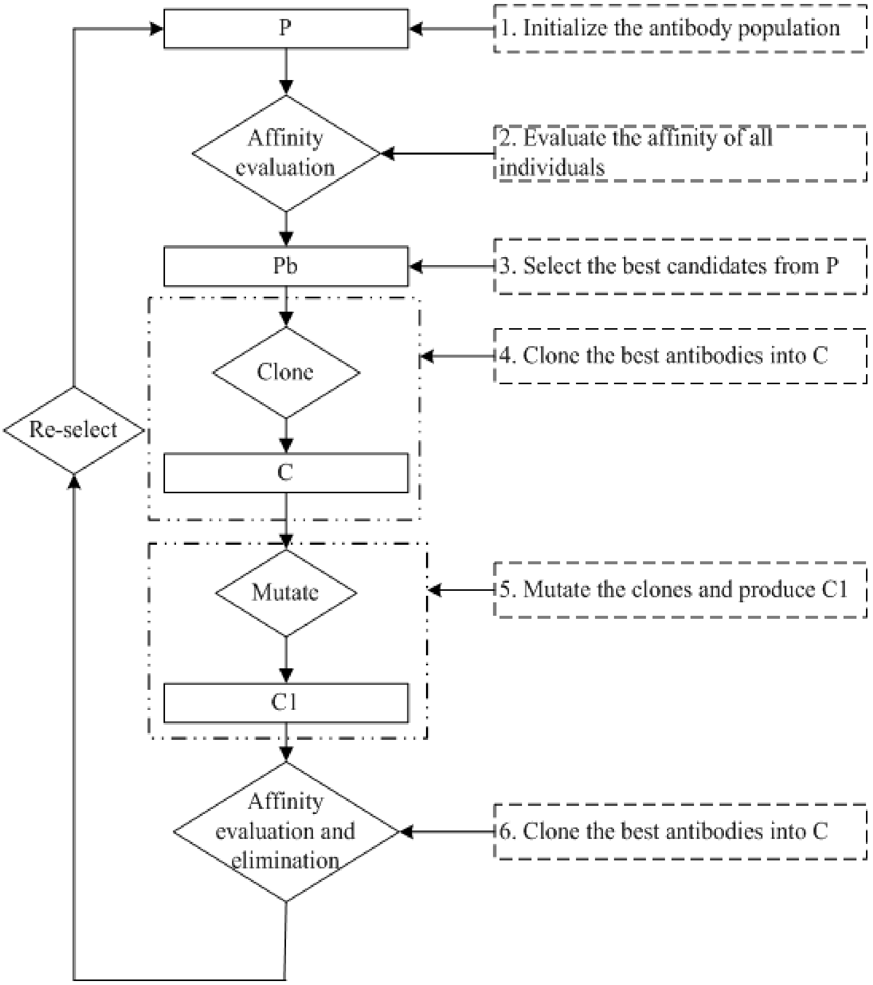

P
1. Initialize the antibody population
Affinity evaluation
2. Evaluate the affinity of all individuals
Pb
3. Select the best candidates from P
Clone
4. Clone the best antibodies into C
Re-select
C
Mutate
5. Mutate the clones and produce C1
C1
Affinity evaluation and elimination
6. Clone the best antibodies into C

Figure 2
Clonal selection method [8]

As its termination condition, CLONALG uses a fixed number of generations, while CLIGA uses the maximum information gain (*K*) principle - $dK/dt \geq 0$. When $dK/dt = 0$, the learning process ends. For t = 0, the following values are computed: fitness value ($P^{(t)}$), cloning expansion (P^{clo}), variation operator ($P^{hyp} = hypermutation(P^{clo})$), evaluation of fitness value of P^{hyp}, the impact of aging ($P^{(t+1)} = aging(P^{hyp}, P(t), \tau_b)$, information gain ($K(t, t_0)$). At the end, the time is increased ($t = t + 1$) and the previous values are recalculated.

B cells distribution function at time *t* is $f^{(t),m}$, where *m* is the fitness value. Information gain can be defined as:

$$\boldsymbol{K(t, t_0) = \sum_m f^{(t),m} \log(f^{(t),m}/f^{0,m})} \quad (1)$$

To evaluate the convergence of algorithms in artificial immune systems, several stopping criteria are used:

- Once a preset number of iterations or function evaluations is reached, the iterative process ceases.
- The iterative process stops as soon as the network attains a preset number of antibodies - convergence of population.
- The average distance between the antibodies and the antigens is examined so as to minimize this value. As soon as the average error is greater than a pre-defined threshold, the iterative process is stopped.
- The network should have converged on condition that its average error increases following *k* consecutive iterations.
- If a defined number of cells do not differ from one network suppression to another, the network is taken to have become stable.
- Once the distance function is inside a specific prescribed distance from the optimum, the algorithm is assumed to have converged.
- The maximum of information gain is reached.

3 Continuous Function Optimization

Continuous optimization is a part of the main mathematical domains for a large number of real-world problems. The problem of continuous optimization includes determining the minimum or maximum value of a function of one or many real variables, subject to constraints (these are, in fact, equations or inequalities). In continuous optimization, the variables in the model can assume any real value within a value spectrum. This feature of the variables is contrary to discrete optimization, where certain variables or all of them may be binary (restricted to the values 0 and 1), integer (for which only integer values are allowed), or more abstract objects drawn from sets with finitely numerous elements [9].

There is a vital difference in continuous optimization between those problems with *no constraints* on the variables and problems where there are *constraints* on the variables. *Unconstrained optimization* problems derive directly from countless practical applications; they further appear in the reformulation of *constrained* optimization problems in which the constraints are replaced by a penalty term in the objective function. *Constrained optimization* problems stem from applications where there are explicit constraints on the variables. There are a great number of subfields of constrained optimization, for which, there are specific algorithms [9].

The applied test functions that belong to the problem domain opt-aiNet are:

- **Multi-function:**

  $\boldsymbol{g(x,y) = x \cdot \sin(4\pi x) - y \cdot \sin(4\pi y + \pi) + 1}$, where x,y ϵ [-2, 2] (2)

- **Roots:**

  $\boldsymbol{g(z) = \frac{1}{1+|z^6-1|}}$, z = x + iy, where x,y ϵ [-2, 2] (3)

- **Schaffer's function:**

  $\boldsymbol{g(z) = 0.5 + \frac{sin^2\left(\sqrt{x^2+y^2}\right)-0.5}{(1+0.001(x^2+y^2))}}$, where x,y ϵ [-10, 10] (4)

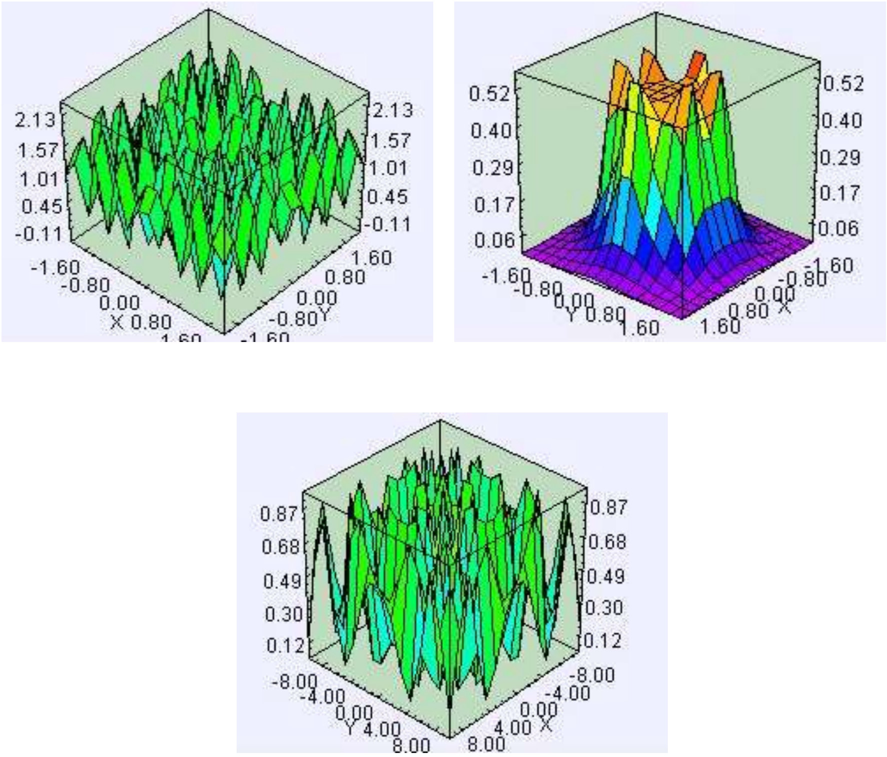


Figure 3

Multi-function, roots, and Schaffer's function

The application of AIS principles to solve different optimization problems is described in [18-21].

4 Case Study - Comparative Analysis of Artificial Immune Algorithms

This paper focuses on examining the performance of artificial immune algorithms in solving continuous function optimization problems using the Optimization Algorithm Toolkit (OAT). OAT is a software environment for developing, evaluating, and experimenting with optimization algorithms on standard benchmark problem domains. There are reference algorithm implementations, graphing, visualizations and various options included in this open-source software. OAT offers a functional library that can be used to investigate existing algorithms and problems, as well as apply new problems and algorithms. This library has a practical explorer and experimenter GUI built on top of it, which ensures that the user can comprehend the functionality in the library. The library's aim is to make easier the best practice of algorithm, problem, and experiment design and implementation, as well as software engineering principles. The GUI supplies a non-technical access that can be used to configure and visualize existing techniques on standard benchmark problem instances [10] [20].

The obtained test results of different artificial immune algorithms are displayed in the following table. The choice of the analyzed algorithms and functions is determined by the built-in capabilities of the OAT software. The run time was set to 180 000 ms (3 minutes) for all analyzed algorithms and the stopping conditions were set at a window width of 1000.

Table 4
Test results of continuous function optimization
(best score/ total evaluations/ iterations for the best score)

	Clonal Selection Algorithm (CLONALG)	**Optimization Immune Algorithm (opt-IA)**	**Optimized Artificial Immune Network (opt-aiNET)**
Multi-function	4.253888443317228/ 5 313 667/22 400	4.253888443317227/ 2 587 598/213 400	4.25388772262681/ 68 334 152/691 000
Roots	0.9999999999999996/ 5 736 550/26 100	1.0/2 517 050/109 000	0.99947323205505/ 28 747 191/1 272 250
Schaffer's function	1.0/5 261 051/21 400	1.0/2 444 295/200 400	0.9999999960496306 /118 526 809/348 900

	Simple Immune Algorithm (SIA)	**CLIGA**	**Optimization Immune Algorithm (opt-IMMALG)**
Multi-function	4.253888443317228/ 7 788 757/15 000	4.2535962571955395/ 5 141 229/4	3.12639125091429/ 71 998 500/919 000
Roots	0.9999999999999996/ 7 764 594/12 700	0.9925992949003002/ 4 761 841/4	0.7162300584615976 /92 545 300/922 300
Schaffer's function	0.9902840901224856/ 7 235 321/18 650	0.9999638750655384/ 5 399 009/1	0.9804511919171228 /86 835 900/1 018 700

By analyzing the results above, it can be concluded that the CLIGA algorithm achieved significantly lower number of required iterations (the third value in the table) for the best score, thus showing the fastest convergence. In addition, having in mind the total number of iterations (the second value in the table), the best result was achieved in the case of the opt-IA algorithm, followed by CLIGA and CLONALG. CLIGA algorithm uses generational aging and information gain stopping criteria. The gain is the quantity of information that the system has learned in relation to the initial distribution function (the randomly generated initial population). This algorithm uses a cloning operator modeled by a cloning potential without memory cells and an aging phase, a stochastic elimination process governed by an exponential negative law, and Kullback's entropy to measure the information gain discovered during the learning process [11]. The Optimization Immune Algorithm (opt-IMMALG) in the cases of multi-function and roots test functions did not achieve the same best score as other tested algorithms.

Conclusions

The main goal of this work is the examination of the continuous function optimization capabilities of artificial immune systems. The performance of these systems, in the form of comparative analysis, was determined using OAT software.

The case study, focused on artificial immune algorithms, showed that the CLIGA algorithm had by far the fastest convergence, the best score - in terms of the number of required iterations, for the analyzed functions. This finding highlighted that with this algorithm, the speed of achieving the used stopping criterion (based on information gain) was the highest. The next algorithms were SIA and CLONALG, which used a specified number of generations, as their stopping criterion.

Further, based on the test results, it can be concluded that the lowest total number of iterations for the defined run-time was achieved with the opt-IA algorithm, followed by the CLIGA and CLONALG algorithms.

Considering the obtained findings, the directions of future research will focus on the application of the CLIGA algorithm, in targeted areas of machine learning.

References

[1] L. de Castro, J. Timmis: Artificial Immune Systems: A New Computational Intelligence Approach, Springer, pp. 57-58, ISBN 978-1-85233-594-6, 2002

[2] C. Chu, M. Lin, G. Liu, Y. Sung: Application of immune algorithms on solving minimum-cost problem of water distribution network, Mathematical and Computer Modelling, Volume 48, Issues 11-12, pp. 1888-1900, 2008

[3] L. de Castro, F. von Zuben, Artificial Immune Systems: Part II - A Survey of Applications, Technical Report DCA-RT 02/00, 2000

[4] G. Q. López, L. A. Morales, L. F. Niño: Immunological computation, Chapter 23, https://www.ncbi.nlm.nih.gov/books/NBK459484/

[5] V. Cutello, G. Nicosia: An Immunological Approach to Combinatorial Optimization Problems, Proceedings of the 8th Ibero-American Conference on AI: Advances in Artificial Intelligence, Seville, Spain, pp. 361-370, 2002

[6] J. Brownlee: Clonal Selection Algorithms, CIS Technical Report 070209A, 2007

[7] B. Ulutas, S. Kulturel-Konak: A Review of Clonal Selection Algorithm and its Applications, Artificial Intelligence Review, Springer, 2011

[8] I. Aydin, M. Karakose, E. Akin: Chaotic-based hybrid negative selection algorithm and its applications in fault and anomaly detection, Expert Systems with Applications, Volume 37, Issue 7, pp. 5285-5294, 2010

[9] Neos Guide, https://neos-guide.org/content/continuous-optimization

[10] Optimization Algorithm Toolkit, https://www.onworks.net/software/app-optimization-algorithm-toolkit-oat

[11] L. de Castro, F. von Zuben: Recent Developments in Biologically Inspired Computing, Idea Group Publishing, 2005

[12] L. de Castro, J. Timmis: An artificial immune network for multimodal function optimization, Proceedings of the 2002 Congress on Evolutionary Computation, Honolulu, HI, USA, IEEE Computer Society, 2002, pp. 699-704, ISBN: 0-7803-7282-4

[13] J. Timmis, C. Edmonds: A Comment on opt-AINet: An Immune Network Algorithm for Optimisation, Proceedings, Part I, Genetic and Evolutionary Computation Conference (GECCO 2004), Seattle, WA, USA, Germany: Springer, 2004, pp. 308-317

[14] L. de Castro, F. von Zuben: Artificial Immune Systems: Part I – Basic Theory and Applications, TR – DCA 01/99, 1999

[15] L. de Castro, J. Timmis: Artificial Immune Systems: A New Computational Intelligence Approach, Springer, pp. 57-58, ISBN 1-85233-594-7, 9781852335946, 2002

[16] E. Hart, J. Timmis: Application areas of AIS: The past, the present and the future, Journal of Applied Soft Computing, Vol. 8, No. 1, pp. 191-201, 2008

[17] L. de Castro, F. von Zuben: aiNet: An Artificial Immune Network for Data Analysis, Data Mining: A Heuristic Approach, Idea Group Publishing, 2001

[18] L. de Castro, F. von Zuben: Learning and optimization using the clonal selection principle, IEEE Transactions on Evolutionary Computation, 2002 Jun, 6(3), pp. 239-251, ISSN: 1089-778X

[19] V. Cutello, G. Nicosia, M. Pavone, G. Narzisi: Real Coded Clonal Selection Algorithm for Unconstrained Global Numerical Optimization using a Hybrid Inversely Proportional Hypermutation Operator, 21st Annual ACM

Symposium on Applied Computing (SAC), Dijon, France, 2006, pp. 950-954, ACM

[20] P. Čisar, S. Maravić Čisar, B. Markoski: Implementation of Immunological Algorithms in Solving Optimization Problems, Acta Polytechnica Hungarica, Vol. 11, No. 4, 2014, pp. 225-240

[21] K. Đuretec: Artificial immune systems, Project: Algorithms based on evolutionary computation, University of Zagreb, Faculty of Electrical Engineering and Computing, www.zemris.fer.hr/~golub/ga/studenti/projekt2008/ais/umjetni_imunoloski_sustavi.pdf

[22] D. Dragulescu, A. Albu: Medical Prediction Systems, Acta Polytechnica Hungarica, Vol. 4, No. 3, 2007, pp. 89-240

[23] J. Jiang, F. Zhang, K. Demertzis: Detecting Portable Executable Malware by Binary Code Using an Artificial Evolutionary Fuzzy LSTM Immune System, Security and Communication Networks, Vol. 2021, pp. 1, 2021

[24] B. Bejoy, T. Bijees et al.: Artificial immune system based frameworks and its application in cyber immune system: A comprehensive review, Journal of Critical Reviews, Vol. 7, No. 2, pp. 52-560, 2020

[25] M. Tabatabaefar, M. Miriestahbanati and J.-C. Grégoire: Network intrusion detection through artificial immune system, Systems Conference (SysCon) 2017 Annual IEEE International, pp. 1-6, 2017

[26] R. Pump, V. Ahlers and A. Koschel: State of the art in artificial immune-based intrusion detection systems for smart grids, 2018 Second World Conference on Smart Trends in Systems Security and Sustainability (WorldS4), pp. 119-126, 2018

[27] R. Pump, V. Ahlers and A. Koschel: Evaluating Artificial Immune System Algorithms for Intrusion Detection, 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), 2020, pp. 92-97, doi: 10.1109/WorldS450073.2020.9210342

Maximizing the Capacity Utilization of Selective Waste Collection Vehicles

Dr. Ádám Titrik, Prof. Dr. habil István Lakatos

Széchenyi István University, Egyetem tér 1, 9026 Győr, Hungary, e-mail: titrika@ga.sze.hu; lakatos@sze.hu

Abstract: Selective Waste Collection is an essential part of recycling raw materials, in order to protect our environment. The waste collection is carried out by using a seriously polluting vehicle, due to the fact that most of the gathering vehicles are using fossil energy sources, like gasoline. High volume of carcinogenic elements are contained in the emitted exhaust gases. The current waste collection methods are just focusing on the load of the selective waste collective vehicle during a collection route. The goal of this research is to find the best solution to use the full storage capacity of the selective waste collecting vehicle with the lowest volume of residual air due to the effectively compressed waste. The closed and uncompressed PET bottles require the largest volume in the waste collecting vehicle. It is essential to minimize the air in the PET bottles to decrease the volume. Different methods have been examined to increase the density of the selectively collected waste. Statistical data have been used to determine the collecting parameters – nowadays, a 15 t waste collecting vehicle, with a 20 m3 load capacity is only gathering 1-1.5 t PET due to the ineffective use of its load compartment. The application of the method herein, enhances the efficiency of the waste collecting vehicle by gathering 4-10 times more waste than currently used methods.

Keywords: selective waste; waste gathering; capacity utilization

1 Introduction

Currently, collecting and reusing materials is common process, especially in case of metals and plastic. In many countries the selective Waste Collecting Vehicles (garbage disposal truck) are using fossil-based energy sources. The engine of these vehicles emit various side-product deriving from the combustion. The exhaust components can cause cancer, birth defects, or other reproductive damages. In order to minimize this kind of pollution, it is important to increase effective payload of the selective waste collecting vehicles. One possible way is to maximize density of the collected waste. The PET bottle is containing too much residual air after the currently applied pre-compression procedure (applied by the user and the vehicle), highly reducing the effectiveness of the waste gathering.

Nowadays the minimization of the air in the waste are reached in the factory/DEPO with different handling technologies. The novelty of this process is the waste density is maximization in the vehicle during its route and not in the factory/ DEPO. Less air in the waste increases the waste density, enhancing the efficiency of collecting by decreasing the frequency of container unloading.

The main topics of the research are:

- Inspecting collecting habits, parameters
- Evaluating statistical data
- Defining and choosing the possible solution
- Testing the possible solution
- Route planning actions with the new method

2 Inspecting the PET Bottle Handling

Investigation of PET collecting habit was carried out on a sample of 150 people, where the results are showed on the following figures. Figure 1 shows that more than 60% of people who collects selectively the PET bottles (polyethylene terephthalate – used for mineral water and soft drinks) put back the cup to the bottle. Some of the Waste Collection Vehicles cannot compress the PET bottles to the minimum possible volume because of the residual air in the closed PET bottles. These PET bottles are working as a spring after the compressing force is reduced in the vehicle, the compressed bottles are expanding back [1].

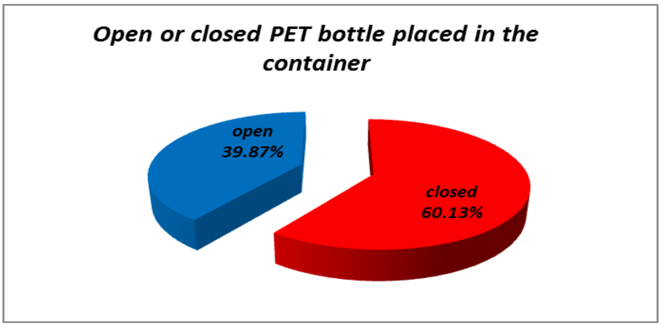


Figure 1
PET bottle handling habits

Volume reduction is applied on the PET bottle with hands predominantly, while 7.33% of the people are not using any compression at all. The distribution of PET bottle volume reduction methods is visualized on Figure 2.

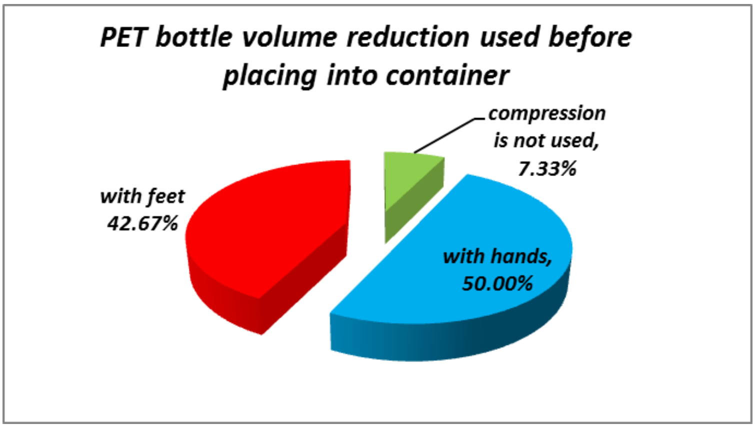


Figure 2
Different volume reduction applied on PET bottle

In Széchenyi István University tests have been made to measure the volume reduction on PET bottle with different type of compression method applied.

Table 1
Achievable volume reduction with different compression method

Bottle volume (PET)	Bottle weight [g]	Original volume [cm^3]	Compression method	Reduced volume [cm^3]	Volume reduction [%]
0.5L	18	520	with hands	470	10
			with feet	390	25
			hand comp.*	440	16
1.0L	28	1030	with hands	540	48
			with feet	276	73
			hand comp.	595	42
1.5L	30	1555	with hands	920	40
			with feet	470	69
			hand comp.	775	51
2.0L	37	2045	with hands	930	54
			with feet	444	78
			hand comp.	930	54
2.5L	45	2540	with hands	1230	52
			with feet	590	77
			hand comp.	960	62

*manual PET bottle compressor

The results clearly show that applying hand operated volume reducer (compressor) does not minimize the PET bottle volume. Nearly the same reduction ratio can be achieved with only hands. By all means, the best volume reduction can be achieved with feet compression (stomping).

Uncompressed and closed PET bottles are the worst for waste collecting vehicles, due to the spring effect of the residual air inside the bottles.

2.1 Collecting Parameters Based on Statistical Data

The statistical data needs to be evaluated for this research, which leads us to define optimized solution for elective waste collection. The examined ~3500 waste container with the capacity of 2.5 m^3, (the data provided by waste handling company, no further parameters can be published) give us the following important data:

- Average filling was 70.4% of the available capacity
- Average selective PET weight was 28.9 kg
- Average density was 11.56 kg/m^3

The fill level and weight of the PET waste is showed in Figure 3, in case of 2.5 m^3 selective waste container (50 pieces, randomly chosen sample from 3500 containers, where the filled level is 100%). The diagram shows that the fulfilled container average weight is ~45 kg. The container max. payload is 900 kg, so the waste density can be increased up to 20 times.

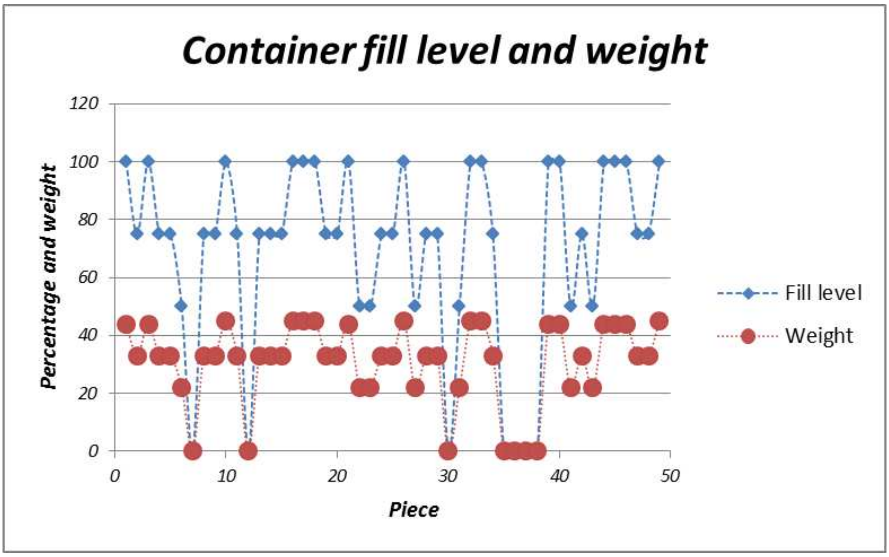


Figure 3

Container fill level and weight relation (PET)

The waste collecting vehicle can gather 35-45 pieces of 2.5 m^3 selective waste container, depending on the collected PET bottle waste density. According to this data, the selective waste collecting vehicle of 20 m^3 capacity can only gather 1000-1300 kg PET waste.

There are two ways for capacity utilization:

- In the container
- In the Waste collection Vehicle

The best solution would be to maximize the waste density at the content of the container. Perforations and volume reduction on the PET bottle can be applied, when it is placed into the container. The applying of shredder is also viable option. Both solutions need external energy to reduce the volume of the PET bottle, however the first solution can use human force through container attached mechanism. It also has to be considered that the usage of human manual labor is trending to minimize. The sufficient energy to operate the container volume reducer is enormous, therefore using a shredder on the vehicle could be a better option.

3 Applying Shredder on PET Bottles

The test was carried out in the Széchenyi István University with the available plastic shredder machine made by F.lli Virginio Srl (Fig. 4). The shredder is not directly suitable for bottles, so the PET bottles have been divided into 3 or 4 pieces for the test.

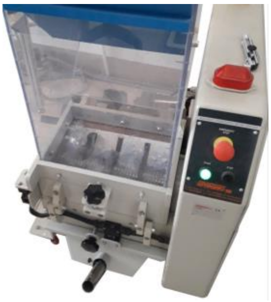


Figure 4
Applied shredder

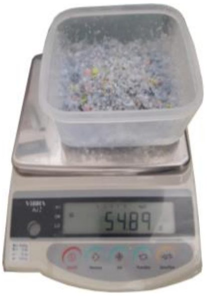


Figure 5
Consistency of the shredded PET bottles

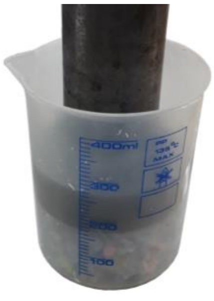


Figure 6
Shredded PET volume reduction test

Test parameters and results

- 3 pieces of 1.5l PET bottles
- Total usage of the shredder was 8 minutes

The applied machine is not optimal for shredding PET bottles. The best shredding performance was not achieved by the used 3 bottles. The placed PET bottles could not be fully shredded because the intake pulling force was decreasing with the size reduction of the bottle pieces. The third of the placed 3 PET bottles stayed in the shredding machine, but the 55 g grinded PET bottle was enough to calculate the result (Fig. 5). The average size of the shredded PET bottles was ~5x5 mm.

The result of the shredding is small plastic pieces, where the size is depending on the parameters of the shredding machine. A measuring bottle was used for analyzing the result, which was filled with shredded plastic pieces to 200 ml without any compression.

Test parameters:

- Waste density for 200 ml shredded PET bottle weights 35.5 g
- Shredded PET waste density is 177.5 kg/m^3
- The shredded PET waste density is 177.5 kg/m^3 compared to the original 11.6 kg/m^3
- Further tests have been made for compressing the shredded waste to 85% (Fig. 6)

Compressing parameters:

- Uncompressed volume: 200 ml
- Compressed volume: 170 ml
- Diameter: Ø72 mm
- Surface: 72*3.14/4=56.52 mm^2
- Applied force: 23 N→ 0.41 N/mm^2

Results in waste vehicle:

- PET waste density 11.6 kg/m^3 →compressed 46.4 kg/m^3 (applying 1:4 volume reduction in waste vehicle)
- PET waste density 177.5 kg/m^3 →compressed 195 kg/m^3 (applying 10% volume reduction in waste vehicle)

According to the results a waste collecting vehicle with 20 m^3 capacity equipped with shredder can gather ~4-10 times more load.

4 Integrated Shredder to Selective Waste Collecting Vehicle

The shredder should be placed on the top of the loading slot of waste collecting vehicle (Fig. 7, red line).

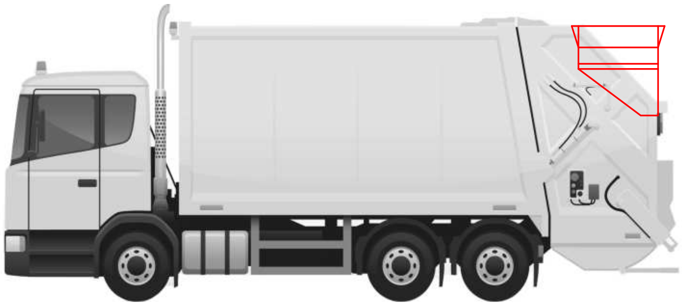


Figure 7
Possible way to attach the shredder to the Waste Vehicle

For the best performance the shredder parameters should be designed to minimize the shredding time and energy consumption. From environmental point of view, the external electrical energy for shredding is better, than using the energy from the internal combustion engine of the vehicle.

4.1. Gathering Plan and Properties of the Shredding

In case of using a shredding unit the route plan will change. In the following section alternative solutions are presented:

Original route plan (Fig. 8):

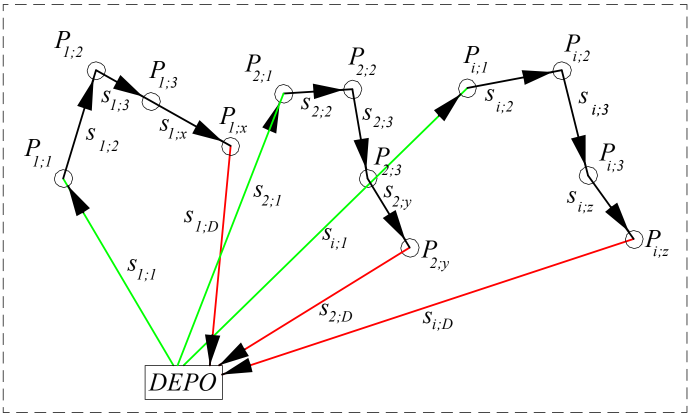


Figure 8
Original gathering plan

Where:

$P_{1;2}$ – P-PET container, $_{1}$- first route, $_{2}$- second container

$s_{1;2}$ – s-distance, $_{1;2}$ from 1st to 2nd containers

$$\Sigma S_{original}=s_{1;1}+s_{1;\,1+i}+s_{1;D}+\ldots+\ s_{i;1}+s_{i;i+1}+s_{i;D}$$

In the original route plan the vehicle collects the container which are alongside the collecting route. The vehicle goes back to the DEPO to empty the waste, when it gets fully filled. After the procedure the vehicle gets back to the next container in the gathering plan. The waste from the container can be compressed up to 25%.

Route plan with shredder equipped vehicle:

In this case the waste vehicle is equipped with shredder, and it can gather more container during its route (Fig. 9).

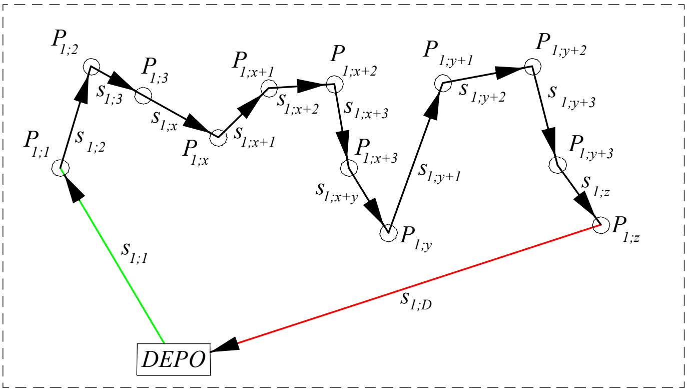


Figure 9
Gathering route with vehicle equipped with shredder

$$\Sigma S_{shredder}=s_{1;1}+s_{1;\,1+i}+s_{1;D}$$

$$\Sigma S_{shredder}<<\Sigma S_{original}$$

The distance of the covered route plan is decreased by using the proposed vehicle shedder method, as the vehicle is able to collect more waste without getting back to the DEPO to empty the collected waste. The traffic load, air and noise pollution are also decreased due to this approach.

Using double sized container at same place for original route plan (Fig. 10):

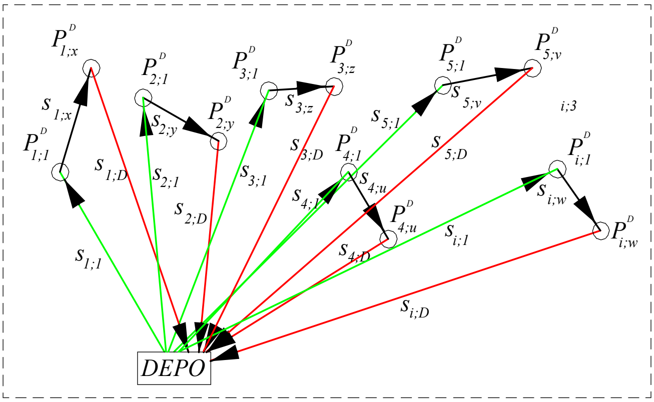


Figure 10

Original gathering plan, double container used

Where:

$P_{D1,2}$ – P-PET container, $_D$- double sized container; $_1$- first route, $_2$- second container

$$\Sigma S_{original,DOUBLE}=s_{1;1}+s_{1;\ 1+i}+s_{1;D}+\ldots+\ s_{i;1}+s_{i;i+1}+s_{i;D}$$

$$\Sigma S_{original,\ DOUBLE}=2*\Sigma S_{original}$$

In this scenario double sized PET container was used in all cases. On the one hand the gathering period duration is decreased by 50%, but on the other hand the waste collecting vehicle is going to be full two times faster. This also doubles the way back to DEPO and to the upcoming container as well.

Gathering route with shredder equipped vehicle, double sized container used:

The waste collecting vehicle can collect 4-10 times more container by using the shredder, although the energy savings can be maximized when double container is used (Fig. 11).

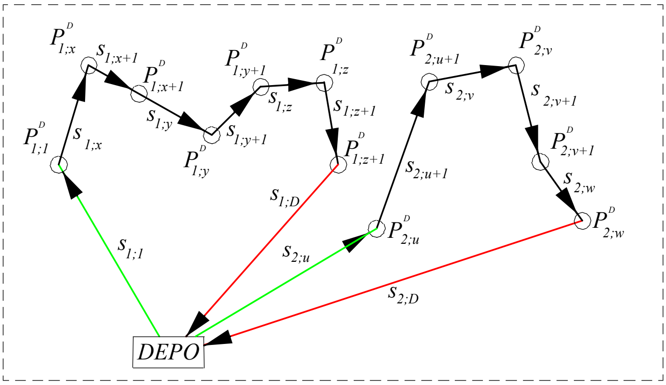


Figure 11
Gathering route with shredder equipped vehicle, double sized container used
$\Sigma S_{\text{shredder, DOUBLE}}=2*\Sigma S_{\text{shredder}}$

In this case the gathering period is decreased by 50%, while the route is getting longer because of the use of shredder (more gathered container). The way back to DEPO and back to upcoming container will be reduced which is both important for environment protection and traffic load reduction.

Conclusions

This work presents a viable solution to maximize the capacity utilization for Waste Collection Vehicles. Reasonable and satisfying result was achieved by the test, where the waste collecting vehicle was able to empty 4-10 times more containers, than in the original gathering solution. The waste density was increased up to 10 times, compared to current traditional methods. In our future work, calculations and optimizations aimed at developing the gathering methods and maximization of energy savings will be researched.

References

[1] Ádám Titrik, István Lakatos; Examining of PET bottle parameters to increase the efficiency of real-time based info-communication waste collection. Budapest, Hungary: Hungarian Academy of Engineering (MMA) (2015) pp. 44-49. Paper: 07, 6 p.

[2] Ádám Titrik, István Lakatos, Dávid Czeglédi: Saturation Optimization of Selective Waste Gathering Vehicle Based on Real–Time Info–Communication System, In: ASME (szerk.) 2015 ASME/IEEE International Conference on Mechatronic and Embedded Systems and Applications. Conference place, date: Boston, USA, 2015.08.02–2015.08.05. New York: American Society of Mechanical Engineers (ASME), 2015. Paper DETC2015–46720. 7 p. (Volume 9) (ISBN:978–0–7918–5719–9) (2015)

[3] Horváth, Adrián ; Hegyi, Csaba: Improving the timing accuracy of route planning methods by developing map databases. LOGISTICS YEARBOOK 2010 pp. 113-121, 9 p. (2010)

[4] Horváth, Adrián: The role of route planning in distribution network decision making. LOGISTICS YEARBOOK 2011 pp. 147-152, 6 p. (2011)

[5] Horváth, Adrián ; Hegyi, Csaba: Simulation analysis of route planning for maximizing the service quality. LOGISTICS YEARBOOK. pp. 123-135, 13 p. (2017)

[6] Horváth, Adrián ; Hegyi, Csaba ; Hirkó, Bálint: Cost analysis of round trips with special regard to pick-up times. LOGISTICS YEARBOOK. pp. 67-83, 17 p. (2018)

[7] Dr. Lakatos, István (2001) Modern emission test of diesel engines in Europe In: Péter, T (szerk.) Symposium on Euroconform Complex Retraining of Specialists in Road Transport, Budapest, Hungary: BME, (2001) pp. 147-153, 7 p.

[8] Bede, Zsuzsanna; Szabó, Géza; Péter, Tamás (2010) Optimalization of road traffic with the applied of reversible direction lanes PERIODICA POLYTECHNICA-TRANSPORTATION ENGINEERING 38 : 1 pp. 3-8, 6 p. (2010)

[9] Péter, T. (2007) Modeling of large nonlinear transport networks. TRANSPORT SCIENCE REVIEW 57 : 9 pp. 322-331,10 p. (2007)

[10] Szauter, Ferenc; Péter, Tamás; Lakatos, István (2014) Examinations of complex traffic dynamic systems and new analysis, modeling and simulation of electrical vehicular systems In: Almas, Shintemirov The 10th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications New York (NY), IEEE/ASME MESA (2014) Paper: 6935613, 5 p.

[11] Adrienn, Buruzs: The environmental impacts of singe use and reusable packaging. In: Özer, Çınar 4th INTERNATIONAL CONFERENCE ON SUSTAINABLE DEVELOPMENT (ICSD): BOOK OF ABSTRACTS Athens, Greece, (2018) p. 89

[12] Buruzs, Adrienn; Torma, András: Reconstruction and Development of Date for Modelling Integrated Waste Management Systems In: Zostautiene, Daiva; Susniene, Dalia; Leisyte, Ludvika 6th International Conference on Changes In Social And Business Environment: CISABE'2016 Bologna, Italy: Medimond Publishing Company, (2016) pp. 1-8, 8 p.

[13] Buruzs, Adrienn; Kóczy, T. László ; Hatwágner, Ferenc Miklós Studies on the sustainability of integrated waste management systems. Proceedings of the 6th Győr Symposium and 3rd Hungarian-Polish and 1st Hungarian-Romanian Joint Conference on Computational Intelligence. Győr, Hungary: Széchenyi István University, (2014) pp. 201-204, 4 p.

[14] Kőrős, Péter; Pusztai, Zoltán: Creating a model for energy-efficient vehicle operation in Simulink. AUTONOMUS VEHICLES-WORKSHOP Győr, Hungary Széchenyi István University (2021) pp. 64-71, 8 p.

Using Machine Learning Algorithms to Detect Malware by Applying Static and Dynamic Analysis Methods

Jakub Palša, Ján Hurtuk, Martin Chovanec, Eva Chovancová

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice
Letná 9, 042 00 Košice, Slovak Republic
email: {jakub.palsa, jan.hurtuk, martin.chovanec, eva.chovancova}@tuke.sk

Abstract: This paper focuses on malware analysis and detection using machine learning methods. The aim of the authors was to perform static and dynamic analysis of programs designed for Windows and then to present the results of the analysis as a dataset. We analysed and implemented different classification methods, such as decision trees, random forests, support vectors and naive Bayes methods. We verified their ability to distinguish malicious and harmless samples and evaluated their success rate using classification accuracy metrics. Then, we compared the results obtained by prediction over the dataset generated by static and dynamic analysis. Classification was more successful on the data gained using the dynamic analysis method. The best malware detection algorithms have been found to be decision tree-based algorithms, in particular the random forest algorithm, which achieves excellent malware detection accuracy of up to 95.95% with a standard deviation of only 0.58%.

Keywords: malware; static analysis; dynamic analysis; dataset; classification

1 Introduction

As the number of every day users using computers/IT systems increases, so does the desire of attackers to obtain and exploit sensitive user information through malware. The number of threats and their severity is constantly increasing, and while in some cases, the damage caused by malware may be imperceptible to the user, in other cases it can lead to severe losses.

As malware evolves over time, the creators of security solutions, aimed to protect systems from malware are seeking and developing new ways to detect it. The traditional method of malware detection using signature recognition is becoming less and less effective as attackers often use various obfuscation techniques to modify malware code to evade this detection method and this method allows only the detection of known malware with known symptoms. Even a minor delay in the response of security solution providers upon the arrival of a new type of malware can cause irreparable damage, which motivates researchers to find more sophisticated ways to detect malicious samples, especially new malicious samples that have not been analysed before.

The drawbacks of traditional malware detection methods are sought to be addressed by machine learning techniques, capable of detecting malware with a high degree of accuracy. Machine learning allows a program to learn from available samples and then to react to new samples, using the learned information. The efficiency of using machine learning to detect malware is boosted by the availability of labeled malicious samples freely available not only to security experts, but now also to the research community. Another factor of this is also the rapid growth of ever cheaper computing power, which allows researchers to speed up machine learning training and to use large quantities of samples.

And that is why in our research we take the path of detecting malicious software through machine learning methods. In the first place in chapter no. 2, we analyse the expertise of researchers who deal with this issue. Next, we describe the sequence of steps of our research. In chapter no. 3 we point out security and its security work with malicious software. Chapter no. 4 describes how we prepared test samples, which were to perform the analysis, which we discuss in Chapter no. 5. In chapter no. 6 we will create a method of creating a data set and in chapter no. 7 we point out the classification methods of machine learning, which were trained, tested and subsequently evaluated on the basis of success. In chapter no. 8 we evaluate and interpret the results of individual machine learning models. In the chapter no. 9 we compare our best results with the results of researchers from chapter no. 2.

2 Related Works

The problem of malware detection using machine learning is not new and has been addressed in many other works. Two key phases have emerged in using machine learning for malware detection: feature extraction from the input data; and selection the most relevant ones that best represent the set of samples and classification. Extracting the features of potentially malicious samples can be done by analysis – static or dynamic [1]. The goal of the analysis is to understand the capabilities of the particular piece of malware, the system parts and files it can attack, its structure, etc. Static analysis is performed without running the analysed sample, as an examination of the structure and code of the analysed sample using various tools [2]. Dynamic analysis focuses on the behaviour of the analysed sample at runtime, observing the interactions with the system and its impact on the system. Both types of analysis have their advantages and limitations and complement each other. After feature extraction, each sample is represented as a feature vector, used by a classification algorithm to train a machine learning model.

2.1 Experimental Results

This section provides an overview of several studies in malware detection and also describes some shortcomings of each approach.

Bai et al. [3] focused on malware detection in a dataset of 19113 executables – 8592 harmless and 10521 malicious samples. As features, they used the information obtained by static analysis of the headers of the executable files. They claimed to have found a total of 197 features, allowing them to distinguish harmless samples from harmful ones; they also used filtering and wrapper methods to select the most appropriate features. In their study, they evaluated the use of classification algorithms – J48 (decision trees) and random forests. To improve the performance of the J48 algorithm, they used combinations of multiple trained models in bagging and boosting techniques. The aforementioned authors performed a total of 3 experiments, differing in the way the data was classified and the choice of features. They concluded that this approach could detect unknown malware with a high accuracy, while maintaining a low false positive rate. The detection accuracies achieved in the experiments ranged from 94.6% to 99.1%. As they concluded, the random forests algorithm and the bagging and boosting techniques significantly increased the classification accuracy, compared to the case when they used the J48 algorithm without using these techniques.

The comparison of various classification algorithms was also discussed by Kumar et al. in [4]. They compared decision trees, random forests, K-nearest neighbors, logistic regression, linear discriminant analysis and naive Bayes algorithms on a dataset of 5210 (2488 harmless and 2722 harmful) samples. As features (68), they used the information obtained from the headers of the analysed executable files. The aforementioned authors achieved the best overall accuracy (98.78%) using the Random Forests algorithm. The worst detection success rate (56.04%) was found when using the Naive Bayes algorithm. An interesting feature of the work was the use of the so-called *integrated feature set*, which was used to increase the overall accuracy of the classification algorithms.

In [5], Moser et al. point out the problems of static analysis in malware detection. They demonstrate obfuscation techniques and point out that static analysis alone may no longer be sufficient for malware detection. In their paper, they conclude that dynamic analysis should be a necessary complement to static analysis, as it is significantly less vulnerable to obfuscating code transformations.

In [6], Firdausi et al. used the K-nearest neighbours, Naive Bayes, Support Vector Machine, J48, and Multilayer Perceptron algorithms on the features obtained by dynamic analysis using Anubis, a freely available dynamic analysis tool. The dataset consisted of executable samples, including a total of 220 unique malware samples. In several different experiments, they achieved the best accuracy (96.8%) using the decision tree-based J48 algorithm. On the contrary, they achieved the worst results (only 62.8%) using the Naive Bayes algorithm. As the authors conclude, malware detection using machine learning combined with dynamic analysis is a fairly effective method.

Shijo & Salim [7] also focused on this approach and took advantage of the benefits of static and dynamic analysis. They used a combination of both types of analysis to detect malicious samples using machine learning on a dataset of 1487 (997 malicious and 490 benign) samples. Static features were strings extracted from executables. They performed dynamic analysis using the Cuckoo Sandbox tool in a secure, virtual environment, outputting a report on the execution behaviour of each sample, listing API calls and registry changes. The authors used 2 algorithms in their work, namely the support vector method and the random forest algorithm. As the authors of the aforementioned paper reported, they obtained best results using the support vector method, namely a detection accuracy of 95.88% for static analysis, 97.1% for dynamic analysis and 98.7% by combining the two. Thus, the achieved results showed that the combination of static and dynamic analysis increased the detection accuracy compared to the use of static and dynamic analysis alone. However, a disadvantage of the study was the smaller number of samples used for training.

2.2 Evaluation of the Experiments

The related works and existing solutions make it clear that using machine learning to detect malicious samples is advantageous and brings a number of benefits over the traditional malware detection approaches. Using a combination of static and dynamic analysis to extract symptoms from individual samples seems to be the most advantageous, as using only one of the methods is no longer sufficient. Research shows that when selecting an appropriate machine learning algorithm, using decision trees to detect malware seems to be a suitable approach – due to the accuracy of detecting malicious samples using decision-tree-based algorithms. Detection accuracy can also be increased by combining multiple trained models, as it is evident in the case of using the Random Forests algorithm. Knowing this, one may design a system capable of classifying a sample as harmful or harmless with high accuracy, based on performing static and dynamic analysis of the particular sample.

Based on the data obtained from previous research experiments, we performed further research. This focused on the combination of static and dynamic analysis and on the combination of several trained malware detection models.

3 Secure Test Environment

An important element of static and dynamic software analysis is the environment, in which the analysis itself takes place. The goal is to create an environment providing no obstacles to the particular piece of malware, allowing its observation in its full beauty. However, it is necessary to prevent malware from breaking out from this environment and causing real damage.

3.1 The Virtual System

For our research, we chose to use *Oracle VirtualBox 6.1* virtualization software. It should be noted that keeping virtualization software up-to-date is key, as many types of malwares attempt to detect execution in a virtual environment and exploit its security flaws to infect the host system.

As the guest virtual operating system, we chose *Windows 7*. At the time, this version of the operating system was widely used and widely deployed. As a result, a large amount of malware targeting this system appeared. Compared to Windows 10, Windows 7 can be modified to execute malicious code more easily, as Windows 10 incorporates a number of automated security features that are laborious to disable and keep disabled.

We installed *Dependency Walker* (a static analysis tool) and *Cuckoo Sandbox* (a dynamic analysis tool, necessary for the execution and to uncover the intent of the particular piece of malware) into the virtual environment. The installation of third-party software also helped to reduce the sterility of the operating system. The latter could cause the malware to detect the execution of the virtual environment and lead to a failure of the analysis.

In order to pretend that the environment is that of a device used daily, for some time, we used the virtual operating system to perform common activities such as browsing web pages, downloading documents from the Internet, playing audiovisual media, etc. This regular use of the system led to the creation of temporary files and registry entries, which also help to mask the fact that it is a virtual system.

To increase the likelihood of successful malware execution, we used older versions of the respective programs.

A very important step in the preparation of the virtual test environment was the modification of the security settings of the Windows operating system. The modifications consisted of disabling the following:

1) *Windows Defender*, a security program that would actively prevent malware execution,
2) the *Windows Update* service, which could install security patches and passively prevent malware from being executed; and
3) *Windows Firewall*, a security program that would monitor the flow of data between networks.

A snapshot of the system was taken after the system configuration was completed. This provides continuous access to the desired virtual operating system configuration, which can be restored any time, preventing lengthy reconfiguration.

Creating a system snapshot is a very important step in performing dynamic analysis. Restoring it allows to negate any impact of the analysed code on the system. It also ensures the same starting conditions for the analysis of each sample. The system snapshot is an important element, also used by the *Cuckoo Sandbox* tool when automating the analysis.

3.2 The Host System

As the host operating system of the workstation, we used *Ubuntu 18.04*. This Linux-based system was chosen because *Cuckoo Sandbox* works best on Linux-based systems. Version 18.04 was necessary because it is the last version of the Ubuntu operating system that both natively supports and includes the *Python 2.7* programming environment. Python 2.7 is required to properly install Cuckoo Sandbox software and its supporting programs, as currently, newer versions (3.x) are not supported by the Cuckoo Sandbox project. The employed version of Cuckoo Sandbox was version 2.0.7.

For added security when working with malware, *virtuaenv*, a Python virtual environment has been established on the host system, without administrative (*sudo*) privileges. With this, every time Cuckoo Sandbox needed to perform an operation requiring such a privilege, the user had to confirm the operation.

4 Preparation of Test Samples

For the purposes hereof, malware samples were obtained from *virusshare.com*, an online malware sample repository [9]. This provides real malware samples for people such as security researchers, forensic analysts, etc. It is maintained by the users themselves, contributing verified malware samples to it.

We downloaded the *VirusShare_00164.zip* package. We chose to use it for its relatively small size (11.88 GB), compared to other packages. Moreover, the more significant reason for its selection was the date it was added – 15 September 2015. This paper focuses on the analysis of malware infecting Windows devices, as this system is the target for the largest amount of available malware.

For this reason, 15 September 2015 is potentially the most appropriate date, as:

- the most widely used version of the Windows operating system in 2015 was Windows 7, with a 62.31% share [10] of all Windows versions.
- Windows 10 was released on 29 July 2015. Thus, back then, a large amount of malware was uploaded to VirusShare.com by the users of Windows 7, the target system for this work.

The healthy samples used in this work are executable programs such as web browsers, audio and video players, UI customization tools, etc. These were obtained from *portablefreeware.com* [11] and *portableapps.com* [12], hosting a large number of downloadable executables.

The downloaded malware sample package in the zip archive contains 65536 malware samples. For the purposes hereof, 3000 executable samples with the .exe extension were selected. Healthy samples are represented by 838 executable programs. Thus, there were approximately 3.6 malware samples for each healthy sample.

A total of 3838 samples were analysed – see Table 1. These were analysed to create the dataset needed for the machine learning process.

Table 1
Number of Samples Prepared for Analysis

sample class	sample count
malware	3000
healthy	838
malware + healthy	**3838**

5 Analysis Execution

After successfully preparing the test samples, we produced the final dataset, which we then used to perform static and dynamic analysis.

5.1 Comparison of Static and Dynamic Analysis

Unlike static analysis, dynamic analysis does not require malware source code, as it can be performed on any application [13] [14]. Unlike static analysis, dynamic analysis can track the actual malware functionality [15], since certain parts of the code, such as an imported library, do not mean active execution of particular library functions.

However, static analysis has several advantages over dynamic analysis (where the examined sample is executed), the most significant of which being speed, security and low requirements [16].

The fact that static analysis does not monitor the behaviour of the programs, comes with certain disadvantages, which are, on the other hand, the advantages of dynamic analysis [17] [18]:

- it is impossible to observe the real behaviour of the particular program;
- it is hard to detect functions actually used;
- it is hard to classify programs with hardly accessible code;
- it is impossible to identify unknown malware.

5.2 Static Analysis

To generate the outputs of the static analysis, we used Microsoft's *Dependency Walker* tool to analyse executable files. It displays all the modules of the monitored file in a hierarchical tree structure. It is freely available for 32 and 64-bit Windows systems.

Using *Dependency Walker*, we analysed the file headers and functions. Then, we saved the obtained information in a text file. Given that over 3000 samples had to be analysed, doing this manually was not an option. Therefore, we used the *Robotask* tool to run *Dependency Walker* and then send instructions to it. It looped through all the samples in the folder and sent the following instructions:

- **CTRL+O** – open dialog box to open the file;
- **absolute path** – the path to the file to be analysed;
- **ENTER** – confirm the selected file.

Then, after the analysis, Robotask sent further instructions:

- **CTRL + S** – save the retrieved files;
- **absolute path** – where to save the data;
- **3 x TAB** – select the format of the file to be saved – we chose „Text with list of imported/exported functions"
- **ENTER** – confirm saving;

The obtained text files contained a huge amount of sample data. For the purposes hereof, it sufficed to focus on the essential details to distinguish a malicious sample from a clean sample. The publicly available Windows API puts enormous power in the hands of malware creators [19], and this is what our research focuses on. In [20], the authors describe the features most commonly used by malware.

5.3 Dynamic Analysis

The first step of dynamic analysis was to configure the *Cuckoo Sandbo*x environment. The option to create a memory dump was disabled after the sample analysis was completed. Creating memory dumps of the virtual operating system for each analysed sample would quickly fill the storage space of the workstation used to perform the analysis. Moreover, this analysis information is not relevant to the purposes hereof.

The analysis mode was set to *headless*, so the analysis would be performed in the background, without allowing on-screen observation of the behaviour of the virtual system. This saved system resources. Moreover, with such a large amount of analysis it would have been impossible to monitor the graphical output of the virtual system. For the same reason, we also disabled the virtual system tool producing screenshots upon any change on the screen during the analysis.

We enabled the possibility to simulate mouse movements in the virtual system during the analysis, to create a more credible impression of the real system running the malicious code.

After finishing the configuration of the *Cuckoo Sandbox* tool, we could actually execute the analysis. The analysis can be initialized either using a command line interface or through a graphical web interface using the *localhost* server on port 8080. Since the web interface was more user-friendly, we chose this for batch execution of the analyses.

The following items can be configured before running the analysis:

- *Network routing* how the sample to be analysed will access the Internet.
- *Package* the file type according to which the appropriate analysis procedure will be selected.
- *Priority* the priority of the analysis of the sample.
- *Timeout* an important parameter of the analysis – this determines how long the analysis will run (in seconds).

6 Creating the Dataset

The dataset is a set of data used to train, test or otherwise work with algorithms. However, it has to have the appropriate form to allow any further operation.

6.1 Processing the Outputs of Static Analysis

After analysing all datasets, it was necessary to consolidate the results into the final dataset form. As the dataset we used the occurrences of importing the 112 selected functions.

Then, we saved the dataset in *comma-separated values (csv)* format. For this purpose, we wrote a Python script to sequentially scan through the obtained text files containing information about the samples, to find the aforementioned features. We were only interested in those occurrences where the function was actually used, i.e. we only searched for imported functions. The search results were written to another text file. At the beginning of this file, there was a header with the "TARGET" entries (i.e. whether the sample was malicious or not), the "filename" (name of the sample) and then the names of all selected functions. Two different versions were created. The first contained the number of occurrences of each of the selected functions, while the second contained only binary information about whether the function was used at least once. One line was created for each sample, with the following format: **clean file – 0, malicious file – 1**, filename, then the number of calls and/or binary information for each selected function. This data was separated by commas, as it is common for *.cvs* files, well-known by the libraries used in machine learning.

After performing the static analysis and converting the results to csv format, the dataset was ready for use. In its final form, it had 3584 records. Its layout is shown in Table 2.

Table 2
Structure of the Dataset after Static Analysis

sample class	sample count
malware	2747
healthy	837
malware + healthy	**3584**

6.2 Processing the Outputs of Dynamic Analysis

After successful analysis of the malware sample by the Cuckoo Sandbox automation software, an analysis report is generated. This provides a summary of the results of all the processes that were performed on the sample. To create a dataset to train and test the machine learning model, all obtained reports have to be processed. To process the data and create the dataset, we used the Python programming language, specifically version 2.7.

The standard data package obtained after analysing the sample is a directory with many subdirectories. However, we were only interested in the *reports* subdirectory, containing the resulting analysis report (called *report*, stored in *json* format).

In order to work with the analysis reports as efficiently and quickly as possible, a script was created – this goes through all the files and deletes the ones that are not named *report.json*. The unnecessary files include various temporary files created by *Cuckoo Sandbox*. This reduces the number of files scanned during the later operations and increases the speed of those operations.

After cleaning-up the working directory containing the analysis reports, we could start processing the reports themselves. The processing consisted of the following steps:

1) List the names of all functions called from the *Windows API* library. The function names will form the attributes of the respective samples. A separate script was created to retrieve all unique function names from all messages.
2) Create the dataset structure. The number of analysed samples indicates the number of rows in the dataset. The header consists of attributes whose count is equal to the number of unique system calls obtained in the previous step.
3) Rescan all reports. The first pass was necessary to determine the number of samples and their total number of attributes to create the dataset structure. The second pass is handled by a similar script, though this time the script adds each function call found in the message to the appropriate column labelled with the name of that function for each single sample found in the dataset.
4) Add a binary identifier as the first attribute of the sample to indicate whether it is malicious or benign. Malware samples had the malware attribute set to 1, while healthy samples had this attribute set to 0.
5) Save the dataset to a file in comma-separated-values (*csv*) format.

The structure of the dataset gained by processing the reports is shown in Table 3.

Table 3

Structure of the Dataset after Dynamic Analysis

sample class	sample count
malware	2937
healthy	828
malware + healthy	**3765**

7 Machine Learning Classification Methods

In this work, four supervised machine learning classification methods ([21], [22], [23], [24]) have been investigated:

1) the Decision Tree method,
2) the Random Forest method,
3) the Support Vector method, and
4) the Naive Bayes method.

For each classification method, a classifier was selected from the *Scikit-Learn* library [25], this was then trained and tested. The data were normalized by scaling using the *StandardScaler()* function. For the support vector method classifier, unlike the other classification methods used, up to 3 models were created, depending on the type of kernel used.

A custom function with nested *for ()* loops was used to tune the hyperparameters, where all combinations of hyperparameter values were tested. Instead of cross-validation, a custom implementation was created. In this, all combinations of hyperparameters were tested 30 times, but at each of the 30 iterations, the dataset was re-segmented into training and test datasets in order to obtain diverse input data. This ensured that the success of the classifier in prediction was verified by using a particular combination of its hyperparameters.

8 Evaluation and Interpretation of Results

The evaluation phase of the prediction model is where the ability of the applied algorithm to correctly classify a sample from the test dataset is verified. Unlike in the learning phase (where, in addition to the training data, the algorithm uses also the attributes of classes of these data), in this phase, when working with the test data it has no information about what group the sample belongs to. In this work, all classifiers were trained on 75% of the input dataset and tested on the remaining 25%. The algorithm assigns a class attribute to the test data – according to its best knowledge – and then its performance is evaluated by the evaluation metric used.

Many evaluation metrics will use *true positive* (TP), *false positive* (FP), *true negative* (TN) and *false negative* (FN) values in their calculations. These values are expressed in a confusion matrix.

The confusion matrix, as shown in Table 4, divides the data into four groups according to their actual and predicted class:

- *true positive* – correctly classified positive samples, i.e. samples correctly classified as malware,

- *false positives* – misclassified negative samples, i.e. harmless samples classified as malware,
- *true negative* – correctly classified negative samples, i.e. samples correctly classified as harmless,
- *false negative* – misclassified positive samples, i.e. malware samples predicted to be harmless.

Table 4
Confusion Matrix

		PREDICTED CLASS	
		positive	*negative*
TRUE CLASS	*positive*	**true positive**	**true negative**
	negative	**false positive**	**false negative**

It is very important to choose an appropriate evaluation metric because not every metric is suitable for all cases. It depends on whether we desire to achieve a high overall success rate for the delivered data or whether the focus is on a particular class. For the purposes hereof, we used the following evaluation metrics [26]:

- classification accuracy and
- sensitivity.

Classification accuracy

The metric of classification accuracy as shown in Equation 1, indicates the proportion of correctly classified samples to all samples. In measuring classification accuracy, the size of the classes is not taken into account and hence no weights are assigned to the classes.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{1}$$

Sensitivity

Sensitivity as shown in Equation 2 is the ratio of correctly classified positive samples to the total number of positive samples. It represents the percentage of correctly identified malware files.

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

8.1 Decision Tree

Table 4 shows the results obtained using the decision tree algorithm. Using static analysis, the decision tree method achieved the highest classification accuracy values of almost 90%. When using dynamic analysis, the values exceeded 94%.

Table 4

Success Rate of the Decision Tree Model

Static analysis			Dynamic analysis		
Parameters	**Accuracy (%)**	**Sensitivity (%)**	**Parameters**	**Accuracy (%)**	**Sensitivity (%)**
DT_S_1	89,74 ± 0,98	95,76 ± 1,08	DT_D_1	94,53 ± 0,74	96,37 ± 0,73
DT_S_2	89,70 ± 0,99	95,75 ± 1,03	DT_D_2	94,38 ± 0,65	96,27 ± 0,69
DT_S_3	89,63 ± 1,11	95,73 ± 1,22	DT_D_3	94,30 ± 0,80	96,10 ± 0,93
DT_S_4	89,61 ± 0,98	95,06 ± 0,77	DT_D_4	94,22 ± 0,78	96,22 ± 0,79
DT_S_5	89,59 ± 1,18	95,10 ± 1,07	DT_D_5	94,18 ± 0,73	96,23 ± 0,71

8.2 Random Forest Method

Table 5 shows the prediction success rate of the Random Forest method. The latter achieved the highest classification accuracy values for both static and dynamic analysis, reaching values exceeding 91% for static analysis and almost 96% for dynamic analysis. This proves that the composite random forest method is more efficient than the decision tree method alone.

Table 5

Success Rate of the Random Forest Model

Static analysis			Dynamic analysis		
Parameters	**Accuracy (%)**	**Sensitivity (%)**	**Parameters**	**Accuracy (%)**	**Sensitivity (%)**
RF_S_1	91,32 ± 0,92	96,94 ± 0,60	RF_D_1	95,95 ± 0,58	98,08 ± 0,49
RF_S_2	91,26 ± 0,88	96,91 ± 0,54	RF_D_2	95,93 ± 0,62	98,06 ± 0,51
RF_S_3	91,25 ± 0,90	97,00 ± 0,61	RF_D_3	95,91 ± 0,70	98,12 ± 0,53
RF_S_4	91,24 ± 0,85	96,84 ± 0,61	RF_D_4	95,90 ± 0,68	98,10 ± 0,53
RF_S_5	91,23 ± 0,88	96,80 ± 0,63	RF_D_5	95,88 ± 0,72	98,10 ± 0,56

8.3 Support Vector Method with a Linear Kernel

The Support Vector method achieved the highest classification accuracy amounting to 87.94% when using static analysis and 95.95% when using dynamic analysis, as shown in Table 6. In dynamic analysis using the Support Vector method, the linear kernel yielded the highest value of classification accuracy of all kernels.

Table 6

Success Rate of the Support Vector Model Using a Linear Kernel

Static analysis			Dynamic analysis		
Parameters	**Accuracy (%)**	**Sensitivity (%)**	**Parameters**	**Accuracy (%)**	**Sensitivity (%)**
SVCL_S_1	87,94 ± 1,02	95,75 ± 0,86	SVCL _D_1	92,38 ± 0,83	97,82 ± 0,62
SVCL_S_2	87,92 ± 0,84	95,44 ± 0,87	SVCL _D_2	92,37 ± 0,84	97,83 ± 0,63
SVCL_S_3	87,01 ± 1,19	96,76 ± 0,70	SVCL _D_3	92,36 ± 0,82	97,81 ± 0,62
SVCL_S_4	84,10 ± 1,10	88,42 ± 1,32	SVCL _D_4	92,25 ± 0,77	97,52 ± 0,63
SVCL_S_5	83,93 ± 1,09	88,22 ± 1,49	SVCL _D_5	92,25 ± 0,75	97,49 ± 0,62

8.4 Support Vector Method with a Radial Kernel

The values obtained using the Support Vector method and a radial kernel are shown in Table 7. The highest classification accuracy was 87.94% in case of static analysis and 92.38% in case of dynamic analysis.

Table 7
Success Rate of the Support Vector Model Using a Radial (rbf) Kernel

Static analysis			Dynamic analysis		
Parameters	**Accuracy (%)**	**Sensitivity (%)**	**Parameters**	**Accuracy (%)**	**Sensitivity (%)**
SVCR_S_1	88,11 ± 0,61	96,69 ± 0,65	SVCR_D_1	91,93 ± 0,84	98,68 ± 0,50
SVCR_S_2	87,84 ± 0,67	96,61 ± 0,67	SVCR_D_2	91,84 ± 0,90	98,46 ± 0,65
SVCR_S_3	87,76 ± 0,73	96,53 ± 0,73	SVCR_D_3	91,71 ± 0,73	98,61 ± 0,57
SVCR_S_4	87,68 ± 0,78	96,94 ± 0,63	SVCR_D_4	91,66 ± 0,82	98,90 ± 0,55
SVCR_S_5	87,60 ± 0,75	96,39 ± 0,74	SVCR_D_5	91,65 ± 0,82	98,88 ± 0,51

8.5 Support Vector Method with a Polynomial Kernel

As far as static analysis is concerned, the highest achieved classification accuracy of the polynomial function kernel of the support vector method amounted to 88.48%. When using dynamic analysis, this value changed to 92.17%. The results are shown in Table 8.

Table 8
Success Rate of the Support Vector Model Using a Polynomial Kernel

Static analysis			Dynamic analysis		
Parameters	**Accuracy (%)**	**Sensitivity (%)**	**Parameters**	**Accuracy (%)**	**Sensitivity (%)**
SVCP_S_1	88,48 ± 0,78	95,46 ± 0,66	SVCP_D_1	92,17 ± 0,79	97,60 ± 0,69
SVCP_S_2	88,38 ± 0,75	95,31 ± 0,70	SVCP_D_2	92,17 ± 0,80	97,59 ± 0,70
SVCP_S_3	88,33 ± 0,80	95,40 ± 0,81	SVCP_D_3	92,16 ± 0,84	97,33 ± 0,77
SVCP_S_4	88,31 ± 0,93	95,39 ± 0,77	SVCP_D_4	92,15 ± 0,83	97,32 ± 0,76
SVCP_S_5	88,23 ± 0,81	95,16 ± 0,69	SVCP_D_5	92,14 ± 0,76	97,60 ± 0,70

8.6 Naive Bayes Method

The naive Bayes method produced the most drastically different results, comparing static and dynamic analysis. However, neither method was able to correctly predict the occurrence of malware samples, as it is evident in Table 9. Also, in case of both analysis types, this method achieved the largest standard deviation values when using the naive Bayes classifier.

Table 9
Success Rate of the Naive Bayes Method Model

Static analysis			Dynamic analysis		
Parameters	**Accuracy (%)**	**Sensitivity (%)**	**Parameters**	**Accuracy (%)**	**Sensitivity (%)**
NB_S_1	42,27 ± 1,28	26,61 ± 1,66	NB_D_1	59,53 ± 1,76	48,58 ± 2,42
NB_S_2	42,22 ± 1,26	26,55 ± 1,63	NB_D_2	59,12 ± 1,51	48,00 ± 2,10
NB_S_3	42,20 ± 1,28	26,51 ± 1,67	NB_D_3	58,43 ± 1,99	47,27 ± 2,67
NB_S_4	42,08 ± 1,26	26,33 ± 1,64	NB_D_4	54,45 ± 2,38	42,40 ± 3,14
NB_S_5	42,05 ± 1,26	26,29 ± 1,64	NB_D_5	48,58 ± 2,00	34,63 ± 2,57

8.7 Summary

Table 10 compares the highest values achieved for each algorithm. Using both types of analysis (static analysis-Figure 1, dynamic analysis-Figure 2), the random forests method, the decision tree method and the support vector method achieved good results.

On the other hand, the naive Bayesian methods could not cope with the particular problem. The most efficient model herein was the random forests model in dynamic analysis, where it achieved a classification accuracy value of 95.95% with a standard deviation of only 0.58%.

Table 10

Comparison of the Best Predictions of the Algorithms Used

Static analysis			Dynamic analysis		
Algorithm	**Accuracy (%)**	**Sensitivity (%)**	**Algorithm**	**Accuracy (%)**	**Sensitivity (%)**
RF	91,32 ± 0,92	96,94 ± 0,60	RF	95,95 ± 0,58	98,08 ± 0,49
DT	89,74 ± 0,98	95,76 ± 1,08	DT	94,53 ± 0,74	96,37 ± 0,73
SVC_P	88,48 ± 0,78	95,46 ± 0,66	SVC_L	92,38 ± 0,83	97,82 ± 0,62
SVC_R	88,11 ± 0,61	96,69 ± 0,65	SVC_P	92,17 ± 0,79	97,60 ± 0,69
SVC_L	87,94 ± 1,02	95,75 ± 0,86	SVC_R	91,93 ± 0,84	98,68 ± 0,50
NB	42,27 ± 1,28	26,61 ± 1,66	NB	59,53 ± 1,76	48,58 ± 2,42

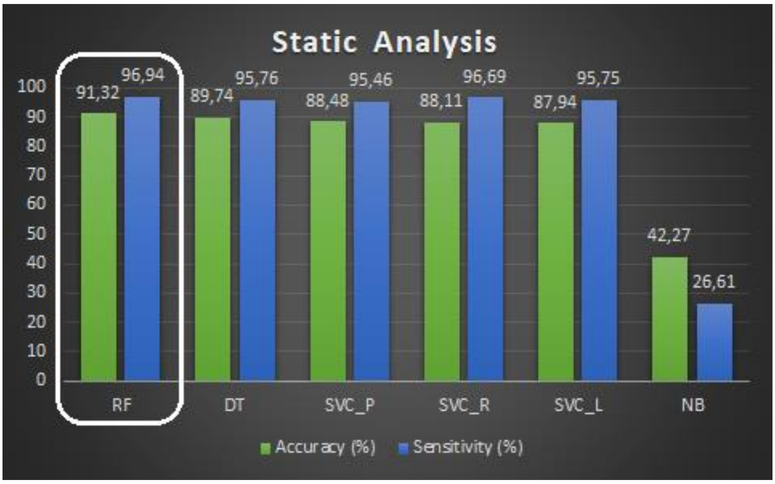


Figure 1

The Highest Values Achieved for Each Algorithm in Static Analysis

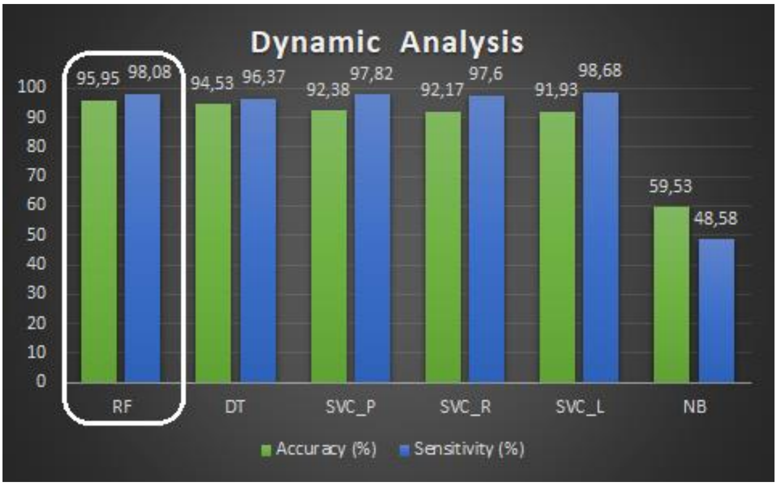


Figure 2
The Highest Values Achieved for Each Algorithm in Dynamic Analysis

9 Comparison of Studies

A comparison of the results of the best algorithms mentioned in the analysed papers and the present study are shown in Table 11. The best malware detection algorithms were found to be decision-tree-based algorithms, especially the random forest algorithm, which achieves excellent malware detection accuracy by aggregating the results of multiple decision tree classifiers for a more accurate result. The algorithm performed well on the features obtained by both static and dynamic analysis, but also in hybrid analysis, where it achieved only 1% less accuracy and sensitivity than the best algorithm, the Support Vector method. The differences in the results of the different papers may be due to the different features and also to their count, as the best accuracies were achieved in the papers sporting fewer features. This was achieved by using methods to select the most relevant flags and removing irrelevant flags that may be useless for the model. The size of the dataset and the different types of malwares in the malicious samples could also have an impact on the results.

Table 11
Comparison of the Best Algorithms from the Analysed Papers

Paper	Analysis	Data	Algorithm	Accuracy (%)	Sensitivity (%)
Bai et al.	static analysis	harmless – 8592 harmful – 10521	Random forests	99,9	99,1
Kumar et al.	static analysis	harmless – 2488 harmful – 2722	Random forests	98,78	99,0
this study	static analysis	harmless – 837 harmful – 2747	Random forests	91,32	96,94
this study	dynamic analysis	harmless – 828 harmful – 2937	Random forests	95,95	98,08
Firdausi et al.	dynamic analysis	harmless – 250 harmful – 220	J48	96,8	95,9
Shijo & Salim	hybrid analysis	harmless – 490 harmful – 997	Support vector method	98,71	98,7

Data comparison from the Table 11 represented in graph Figure 3.

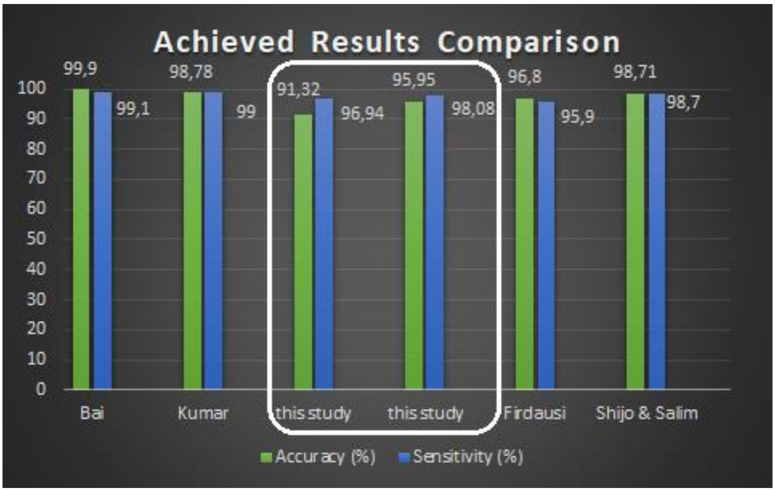


Figure 3
Comparison of the Best Algorithms from the Analysed Papers in Graph Form

Conclusion

Of all machine learning classification methods used, random forests achieved the best results in both types of analyses.

In addition to random forests, decision trees and support vector methods also achieved solid results.

The Naive Bayes methods did not prove to be able to correctly detect malware samples, compared to the other machine learning methods used.

In general, all algorithms achieved higher values of both classification accuracy and sensitivity when using the dataset created by dynamic analysis of the samples. However, by combining static and dynamic analysis and also by combining multiple trained models, the accuracy of malware detection improved.

Acknowledgement

This work was supported by the Faculty of Electrical Engineering and Informatics at the Technical University of Košice under contract No. FEI-2021-76: A modern interpreter of predicate linear logic formulas and by KEGA 002TUKE-4/2021 Implementation of Modern Methods and Education Forms in the Area of Cybersecurity towards Requirements of Labour Market.

References

[1] A. Fedák and J. Stulrajter, Fundamentals of Static Malware Analysis: Principles, Methods, and Tools., 2010, In Science and Military, pp. 45-53

[2] E. Masabo, K. Kaawaase, J. Sansa-Otim, J. Ngubiri, and D. Hanyurwimfura., 2018, In A State of the Art Survey on Polymorphic Malware Analysis and Detection Techniques

[3] J. Bai, J. Wang, and G. Zou., A Malware Detection Scheme Based on Mining Format Information., 2014, In TheScientificWorldJournal

[4] A. Kumar, K. S. Kuppusamy, and A. Gnanasekaran., A learning model to detect maliciousness of portable executable using integrated feature set., 2017, In Journal of King Saud University - Computer and Information Sciences

[5] A. Moser, Ch. Kruegel, and E. Kirda., Limits of Static Analysis for Malware Detection., 2007, In Twenty-Third Annual Computer Security Applications Conference., pp. 421-430

[6] I. Firdausi, Ch. Lim, A. Erwin, and A. Nugroho., Analysis of Machine learning Techniques Used in Behavior-Based Malware Detection., 2010, In Advances in Computing, Control, and Telecommunication Technologies, International Conference., pp. 201-203

[7] P. V. Shijo, and A. Salim., Integrated Static and Dynamic Analysis for Malware Detection., 2015, In Procedia Computer Science

[8] N. Lutsiv, T. Maksymyuk, M. Beshley, O. Lavriv, L. Vokorokos, and J. Gazda., Deep Semisupervised Learning-Based Network Anomaly Detection., 2022, In Heterogeneous Information Systems. CMC-Computers, Materials & Continua., pp. 413-431

[9] Online repository of malware samples VirusShare.com [Online] Avaliable: https://virusshare.com

[10] Statcounter., GlobalStats., [Online] Avaliable: https://gs.statcounter.com/windows-version-market-share/desktop/worldwide/#monthly-201001-202001

[11] PortableFreeware.com., [Online]. Avaliable: https://www.portablefreeware.com

[12] PortableApps.com., [Online]. Avaliable: https://portableapps.com

[13] N. Ádám, B. Madoš, A. Baláž, and T. Pavlik., Artificial neural network based IDS., 2017, In Proc. of the 15th International Symposium on Applied Machine Intelligence and Informatics., pp. 159-164

[14] U. Bayer, E. Kirda, and C. Kruegel, Improving the efficiency of dynamic malware analysis., 2010, In Proceedings of the 2010 ACM Symposium on Applied Computing., pp. 1871-1878

[15] Z. Dankovičová, D. Sovák, P. Drotár, and L. Vokorokos,. Machine learning approach to dysphonia detection.2018 In Applied Sciences., p. 1927

[16] B. Kang, T. Kim, H. Kwon, Y. Choi, and E. G. In., Malware Classification Method via Binary Content Comparison., 2012, In Proceedings of the 2012 ACM Research in Applied Computation Symposium., pp. 312-316

[17] A. Baláž, N. Ádám, E. Pietriková, and B. Madoš., ModSecurity IDMEF module., 2018, In Proc. of the 16th World Symposium on Applied Machine Intelligence and Informatics (SAMI), pp. 77-81

[18] D. Uppal, M. Vishakha, and V. Verma., Basic survey on Malware Analysis, Tools and Techniques., 2014, In International Journal of Computer Science & Applications., pp. 103-112

[19] H. Tamada et al., Dynamic software birthmarks based on API calls., 2006 In IEICE Transactions on Information and Systems

[20] M. Sikorski, and A. Honig., PRACTICAL MALWARE ANALYSIS., 2012, In no starch press

[21] T. K. Ho., Random decision forests. In Proceedings of 3rd International Conference on Document Analysis and Recognition., 1995, In Montreal, Quebec, Canada, pp. 278-282

[22] L. Breiman, and A. Cutler., "Random forests." Machine learning., 2014, In Netherlands, pp. 5-32

[23] H. Zhang., Exploring conditions for the optimality of naive Bayes., 2005, In International Journal of Pattern Recognition and Artificial Intelligence., pp. 183-198

[24] S. Suthaharan., Support vector machine., 2016, In Machine learning models andalgorithms for big data classification., Springer, Boston, pp. 207-235

[25] F. Pedregosa, et al., Scikit-learn: Machine Learning, In {P}ython., 2011, In Journal of Machine Learning Research., Vol. 12, pp. 2825-2830

[26] I. H. Witten, and E. Frank., Data mining practical machine learning., 2005, In Amsterdam: Morgan Kaufman, pp. 53-129

Multi Cantilever-Mass Mechanism for Vibration Suppression

Péter Szuchy[1], Lívia Cveticanin[2], István Bíró[3]

[1,2]Doctoral School on Safety and Security Sciences, Óbuda University, Népszínház u. 8, 1081 Budapest, Hungary, cpinter.livia@bgk.uni-obuda.hu

[1,3]Department of Mechanics, Faculty of Engineering, University of Szeged, Mars tér 7, 6724 Szeged, Hungary, gi@mk.u-szeged.hu

Abstract: In this paper a new type of passive mechanism for vibration suppression is introduced. The mechanism is based on the system of cantilever – mass units (dynamic absorbers) connected to the basic structure. The support of cantilevers is rigid or even elastic. The vibration of the system is caused by external excitation force which acts on the basic structure. The aim of the paper is to determine the parameters of the system for which the frequency gap and the vibration suppression occur. The used mathematical model is a system of coupled equations where the measured parameters are introduced by the application of a newly developed, so-called, 'elastic support method'. Solving the mathematical model, the amplitude-frequency vibration property of the system is obtained. The computed solution is compared with that the previously published result of the 'wallpaper' type metastructure for vibration suppression, which is modeled as a system of translation moving system of mass-in-mass units. It is concluded that the effect of the suggested mechanism is in good agreement with that of the metastructure for vibration suppression. The resonances of the two models are matched with the results of Inventor Finite Element Analysis, too. Difference in results is negligible.

Keywords: "wallpaper"-like metamaterial; cantilever-mass mechanism; vibration suppression; 5-DoF system; natural-frequency

1 Introduction

Recently, mechanical metastructures and metamaterials are developed for suppression or elimination of vibration. Metamaterials and metastructures are artificially composed systems containing a basic mass in which small masses are added [1-4]. The added masses have the role of vibration absorbers. Opposed to the conventional materials, the metastructure absorbers are integrated into the basic material [5]. Metastructures are modeled as complex systems of mass-in-mass units where properties of the added mass-spring unit satisfy the condition for dynamic absorber of the basic mass [6-9]. Various types of metastructures are already developed: in-line 1D or bar structures [10-13], space structures [14-15] and plane or wall panels [16-18]. The main disadvantage of all of these metastructures is the complexity of their fabrication. The new types of metastructures have absorbers made of the same material as the basic structure and the system is constructed as a single unit. The 3D printing technique allows creation of such structures, with extremely complex geometries tuned for broadband vibration suppression, for example a square structure inside of which mass as absorbers act [19-21] or stick – like resonator [22]. Such 3D-printed metastructures are suitable for passive vibration suppression. In addition, the structures remain capable of bearing loads without adding additional mass. For all of the mentioned metamaterials, it is common that they suppress vibration in certain frequency region [23, 24], and the width of the band gap, where the decrease of the amplitude of vibration occurs, is very small.

To eliminate these lacks, the system with the higher number of dynamic absorbers [25] and nonlinear properties [26] are introduced. Thus, the 'wallpaper'-like metastructure which contains 5 different dynamic absorbers is able to admit 5 different vibration frequencies [18]. In spite of the fact that the design seems to be simple (between a basic plane and an external surface with cups masses are settled (see Fig. 1) which move translator up and down due to the action of the vertical external excitation), fabrication with proper values for vibration suppression is not an easy task.

In this paper the new multicantilever-mass mechanism for vibration suppression is developed.

The aim of the new mechanism is to eliminate vibration on certain frequencies, of all the unwilling ones. The requirements for vibration elimination directly influence the design of the mechanism and the number of cantilever-mass dynamic absorbers. In the paper, dynamics of the mechanism is mathematically modeled. Parameters of the model are included by using the new, so-called 'elastic support method', springing from measured values. After solving the equation of the oscillatory motion, the results are applied for parameter analysis of the mechanism. The new mechanism with 5 cantilever-masses is compared with the 5-DoF wallpaper like metastructure. It is observed that the comparison is possible and that the parameters of the new model are more controllable than the previously developed wallpaper like metastructure with the translation motion of masses. The resonance cases of both models are matched using of Inventor Finite Element Analysis. Difference in results is negligible.

The paper has 5 sections. After the introduction in Section 2, the physical and mathematical model of the cantilever-mass mechanism is developed. The system is described with n linear coupled second order differential equations. In Section 3, a new method for calculation of the stiffness of the elastic support is developed. In Section 4, the amplitude-frequency vibration property of the 5-DoF cantilever-mass mechanism is obtained. The result is compared with that obtained for the 5-DoF ‘wallpaper’ type mechanism with translator motion. The solution is matched with the results of Inventor Finite Element Analysis (FEA) of the 3D solid body. The paper ends with conclusions.

2 Model of the Mechanism

In this section, the physical and mathematical model of the cantilever-mass mechanism is developed.

2.1. Physical Model of the Cantilever-Mass Mechanism

Based on the principle of the ‘wallpaper’ metastructure, but keeping the practical needs of measuring in mind, the translational model (Fig. 1) is changed into the cantilever beam model (Fig. 2). The cantilever-mass mechanism for vibration suppression is physically modelled as a system of beams clamped at one end with concentrated masses on the other end, which are attached to the primary structure. The primary structure is modelled as a clamped beam-mass unit (Fig. 2a), where the rigidity of the beam is EI_1 and the mass is m_1. On the mass m_1, the periodic excitation force $F(t)$ acts, which causes vibrations. Each added unit, which represents a dynamic absorber, contains a cantilever of transversal rigidity EI_i, a length b_i, and concentrated mass m_i. Only one damper with coefficient k_1 remained. The number of units is not limited, it depends on the requirement for elimination of vibrations of the basic element 1.

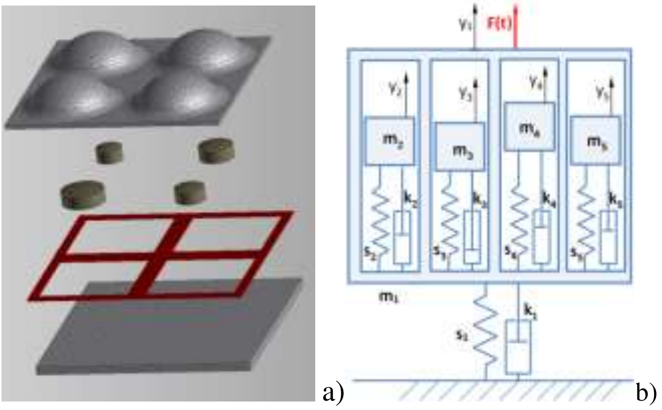


Figure 1

“Wall-paper”-like metastructure: a) Exploded view of the 3D model; b) 5-DoF translational model

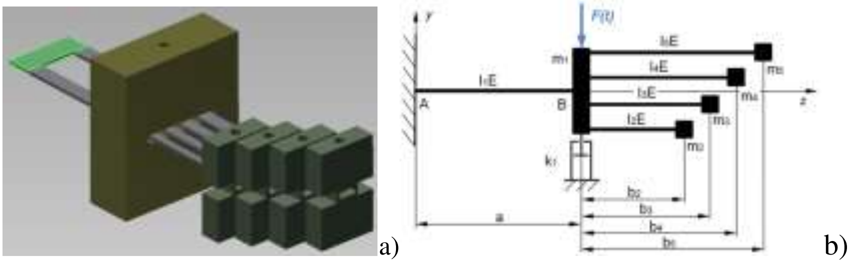


Figure 2
The 5-DoF cantilever-mass mechanism: a) 3D solid body model; b) Scheme of model

2.2. Mathematical Model of the Cantilever-Mass Mechanism

Let us consider the mathematical model of the multicantilever-mass mechanism where on the basic structure 1, the *i*=2,3,…*n* absorber units are attached (see Fig. 2b). It is supposed that the beam bending is in one plane. In addition, the movement of discrete masses is assumed to be in-line.

The bending position of each mass m_j is obtained by applying the Betti theorem summarizing the deflection of the mass m_1 caused by the external force, deflections of m_j due to external and inertial force and the displacement caused by interaction between added masses of the absorber.

Using the linear bending theory, the displacement of the mass m_1 under influence of the force F_1 is obtained as

$$y_B = \frac{F_1 a^3}{3\, I_1 E} \tag{1}$$

where *a* is the position of the mass, i.e. the length of the beam 1. It gives the inverse rigidity coefficient

$$c_{11} = \frac{y_B}{F_1} = \frac{a^3}{3\, I_1 E} \tag{2}$$

However, the force F_1 causes bending of all beams in the system and the bending position of masses m_i are

$$y_{Ci} = \frac{F_1 a^2 (2a + 3b_i)}{6\, I_1 E} \tag{3}$$

The corresponding inverse rigidity coefficients are

$$c_{1i} = \frac{a^2 (2a + 3b_i)}{6\, I_1 E} \tag{4}$$

where *i*=2,3,...*n*.

According to the force F_i, which directly acts on m_i, the displacements are

$$y_{Ci} = \frac{F_i a(a^2+3ab_i+3b_i^2)}{3\,I_1 E} + \frac{F_i b_i^3}{3\,I_i E} \tag{5}$$

and the inverse rigidity coefficients follow as

$$c_{ii} = \frac{a(a^2+3ab_i+3b_i^2)}{3\,I_1 E} + \frac{b_i^3}{3\,I_i E} \tag{6}$$

The force F_i which acts on m_i has also an influence on the other masses m_j of the mechanism. Thus, for i,j=2,3,...n, $i\neq j$ the displacement is

$$y_{Ci} = \frac{F_2 a^3}{3\,I_1 E} + \frac{b_2 F_2 a^2}{2\,I_1 E} + \left(\frac{F_2 a^2}{2\,I_1 E} + \frac{b_2 F_2 a}{I_1 E}\right) b_3 = \frac{F_i a(2a^2+3ab_i+3ab_j+6b_i b_j)}{6\,I_1 E} \tag{7}$$

and the corresponding inverse rigidity coefficient

$$c_{ij} = \frac{a(2a^2+3ab_i+3ab_j+6b_i b_j)}{6\,I_1 E} \tag{8}$$

Remark: There is the symmetry of the rigidity coefficients, and for $i\neq j$ it yields $c_{ij}=c_{ji}$, where i,j=2,3,...n.

Using the previous consideration, the total deflection of each mass (including the mass m_1) is calculated as

$$y_i + \sum_{j=1}^{n} c_{ij} F_j = 0 \qquad i = 2,3 \ldots n \tag{9}$$

Introducing the inertial, damping and excitation force acting on mass m_1,

$$F_1 = m_1 \ddot{y}_1 + k_1 \dot{y}_1 - F_0 \sin \omega_g t \tag{10}$$

and forces acting on masses m_j, where j=2,3,...n

$$F_j = m_j \ddot{y}_j + k_j \dot{y}_j \tag{11}$$

the system of differential equations of motion for the system follows as

$$\boldsymbol{M}\ddot{\boldsymbol{y}} + \boldsymbol{K}\dot{\boldsymbol{y}} + \boldsymbol{C}^{-1}\boldsymbol{y} = \boldsymbol{F} \tag{12}$$

where **C** is the symmetric matrix of the inverse rigidity coefficients, ***K*** is the damping matrix, ***M*** is the mass matrix, i.e.

$$\boldsymbol{C} = \begin{bmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{bmatrix}, \qquad \boldsymbol{K} = \begin{bmatrix} k_1 & 0 & \ldots & 0 \\ 0 & k_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & k_n \end{bmatrix}, \qquad \boldsymbol{M} = \begin{bmatrix} m_1 & 0 & \ldots & 0 \\ 0 & m_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & m_n \end{bmatrix}$$

$$\boldsymbol{y} = [y_1, y_2, \ldots., y_n] \qquad \text{and} \qquad \boldsymbol{F} = [F_0 sin\omega_g t, 0, \ldots., 0]. \tag{13}$$

The equation (12) is a system of *n* coupled second order differential equations which describes the absorber motion. Solving the system (12), the amplitude-frequency relations are obtained.

3 Model of the Beam Fixed in an Elastic Support

The model (12) corresponds to the case when masses of cantilevers are omitted. If the mass of the cantilever is as significant as the value of added mass, it has to be taken into consideration. It is supposed that the mass of the beam is continually distributed along its length, the total mass of the basic beam is m_{r1}, while the masses of the beams in mechanism are m_{ri}, where i=2,3,…n. Reducing masses of beams in the position B of the basic mass (Fig. 3a) the total masses m_{red1} and m_{redi} (i=2,3,…n) are obtained.

Thus, for the system which contains only one added cantilever beside the basic one (Fig. 3a), the reduced mass in B is obtained by equating the kinetic energy of the distributed and point masses, i.e.

$$\frac{1}{2}m_{red1}\dot{y}_B^2 = \frac{1}{2}\frac{m_{r1}}{a}\int_0^a(\dot{y}_1(z))^2\,dz, \qquad \frac{1}{2}m_{red2}\dot{y}_B^2 = \frac{1}{2}\frac{m_{r2}}{b_2}\int_a^{a+b_2}(\dot{y}_2(z))^2\,dz \tag{14}$$

where $\dot{y}_B$ is the velocity of B, $\dot{y}_1(z)$ and $\dot{y}_2(z)$ is velocity distribution along the beam 1 and 2, respectively.

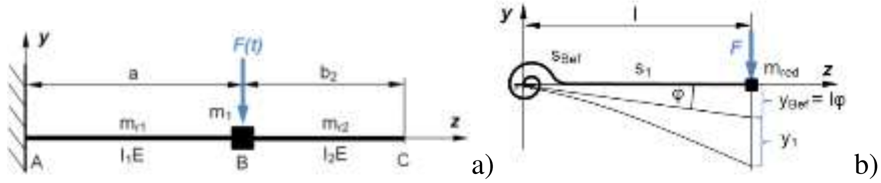


Figure 3

Scheme of the cantilever supported: a) rigid; b) elastic

Using the assumption that there is a direct correlation between the velocity $\dot{y}$ and displacement y i.e. $\frac{\dot{y}(z)}{\dot{y}_B} = \frac{y(z)}{y_B}$ where $\dot{y}_B$ is the velocity and y_B is the displacement of B, equations (14) transform into

$$\frac{1}{2}m_{red1}\dot{y}_B^2 = \frac{1}{2}\frac{m_{r1}}{a}\frac{\dot{y}_B^2}{y_B^2}\int_0^a(y_1(z))^2\,dz, \qquad \frac{1}{2}m_{red2}\dot{y}_B^2 = \frac{1}{2}\frac{m_{r2}}{b_2}\frac{\dot{y}_B^2}{y_B^2}\int_a^{a+b_2}(\dot{y}_2(z))^2\,dz \tag{15}$$

where the deflection of B is $y_B = \frac{Fa^3}{3I_1E}$ and the elastic lines of AB and BC [14]

$$y_1(z)=\frac{F}{6I_1E}(3az^2-z^3),\ \ y_2(z)=\frac{Fa^2}{6I_1E}(3z-a). \tag{16}$$

After integration of (15) with (16) and some modifications, the reduced masses m*red1*and m*red2* are obtained as

$$m_{red1}=\frac{33}{140}m_{r1},\qquad m_{red2}=m_{r2}\frac{3(a+b_2)^2+a^2}{4a^2} \tag{17}$$

Usually, it is assumed that the cantilever is connected with the rigid support. However, in multicantilever-mass mechanism, the units are elastically supported. In the next section, a method is developed for including the effect of elastic support in the rigidity coefficient of the system. The method represents a mixed analytic-experimental one, where the measured vibration values are incorporated into the model of the system.

3.1. Correction of Stiffness Matrix using the 'Elastic Supporting Method'

The procedure for including correction in the coefficient of rigidity of the system due to elastic support is named 'elastic supporting method'. On the system with reduced mass m*red* and a beam, with length l and rigidity s_1, clamped in elastic support with rigidity s_{Bef}, the excitation force F acts (Fig. 3b). It causes the displacement of B, which is the sum of the bending of the beam y_1 and displacement y_{Bef} due to inclination for angle φ. For the inverse rigidity of the beam $c_1=\frac{Fl^2}{3IE}$, the bending displacement is

$$y_1=Fc_1 \tag{18}$$

It is assumed that the displacement of B due to the bending torque Fl is the linear function of the inclination angle φ. For $\varphi=(Fl)c_{Bef}$ and $y_{Bef}=l\varphi$, the displacement is

$$y_{Bef}=Fc_{Bef}l^2 \tag{19}$$

where c_{Bef} is the inverse rigidity coefficient of support and l is the length of the beam. Finally, due to elasticity of the beam and of the elastic connection, the total displacement in B is

$$y=y_{Bef}+y_1=l\varphi+\frac{Fl^3}{3IE}=F(c_{Bef}l^2+c_1) \tag{20}$$

According to (20) the reduced inverse rigidity coefficient follows as

$$c_{red}=\frac{y}{F}=c_{Bef}l^2+c_1 \tag{21}$$

For the reduced inverse rigidity (21) and reduced mass (17), the frequency of vibration is obtained in the form

$$f_m = \frac{1}{2\pi}\sqrt{\frac{1}{c_{red}m_{red}}} = \frac{1}{2\pi}\sqrt{\frac{1}{(c_{Bef}l^2+l^3/3IE)m_{red}}} \qquad (22)$$

where $m_{red} = m_1 + m_{red1} + m_{red2}$

Relation (22) is suitable for determination of the unknown rigidity coefficient c_{Bef}. Measuring the frequency of vibration of the model and substituting the obtained value into (22), the rigidity coefficient of the support c_{Bef} is calculated. The method for obtaining of the unknown rigidity suggested in the paper is a mixed procedure which interacts the data of the measurement and the analytically computed value.

To prove the accuracy of the suggested method, the comparison of calculated rigidity coefficient of a 1-DoF flexible cantilever-mass beam with experimentally obtained one is done (see Fig. 4). Position of the mass in the mechanism is varied and the frequency of the system is changed. In spite of that, it is obtained that both the calculated and the measured inverse rigidity coefficient of support c_{Bef} remain almost constant. Difference between measured and computed values is negligible.

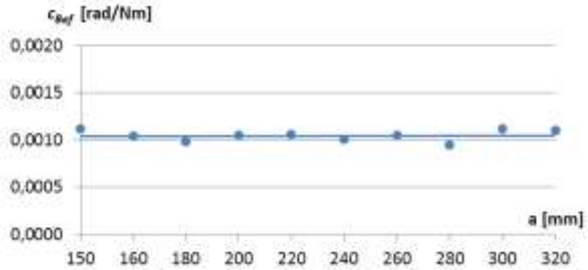


Figure 4

Inverse rigidity coefficient of support (c_{Bef}) of 1-DoF flexible clamped cantilever beam for different positions of mass m_1 (a) obtained by measuring (dots) and by computed trend curve (full line)

Using the suggested procedure, the elements of the symmetric matrix **C,** where the inverse rigidity coefficient of support is included, are calculated as

$$c_{11} = c_b a^2 + \frac{a^3}{3\,I_1E} \qquad (23)$$

$$c_{1i} = c_{i1} = c_b(a^2 + ab_i) + \frac{a^2(2a+3b_i)}{6\,I_1E}, \quad (i = 2 \dots 5) \qquad (24)$$

$$c_{ii} = c_b(a + b_i)^2 + \frac{a(a^2+3ab_i+3b_i^2)}{3\,I_1E} + \frac{b_i^3}{3\,I_iE}, \quad (i = 2 \dots 5) \qquad (25)$$

$$c_{ij} = c_{ji} = c_b\left[a^2 + a\left(b_i + b_j\right) + b_ib_j\right] + \frac{a\left[2a^2+3a(b_i+b_j)+6b_ib_j\right]}{6\,I_1E}, \quad (i,j = 2 \dots 5, i \neq j)\mathrm{c_{ij}} = \mathrm{c_{ji}} = \mathrm{c_b}\left[\mathrm{a}^2 + \mathrm{a}\left(\mathrm{b_i} + \mathrm{b_j}\right) + \mathrm{b_ib_j}\right] + \frac{\mathrm{a}\left[2\mathrm{a}^2+3\mathrm{a}(\mathrm{b_i}+\mathrm{b_j})+6\mathrm{b_ib_j}\right]}{6\,\mathrm{I_1E}}, \quad (\mathrm{i},\mathrm{j} = 2 \dots 5, \mathrm{i} \neq \mathrm{j}) \qquad (26)$$

It is important to emphasize that the effect of the inverse rigidity coefficient of support on the frequency of the system is up to 20%.

During experimental investigation it is seen that the value of the inverse rigidity coefficient of support is influenced by several factors (including the structure of the vise, the supporting force and the material of the support soil).

4 Comparison of Vibration Properties of the Translational and the Cantilever-Mass Models

In our investigation, the 5-DoF cantilever-mass mechanism which contains 4 absorbers settled on the basic mass m_1 is considered (Fig. 2b). Motion of the beams is assumed to be only in one plane and of the end points in vertical direction. Using the AutoCAD Inventor Software the 3D solid body system is created (Fig. 2a). Parameters of the mechanism are: a=0.12 m, b_2=0.11 m, b_3=0.13 m, b_4=0.15 m, b_5=0.17 m, m_1=4.000 kg, m_2=m_3=m_4=m_5=0.500 kg, ρ=7850 kg/m^3; E=210 GPa, I_1E=12.285 Nm^2, I_2E=2.835 Nm^2, I_3E=3.78 Nm^2, I_4E=4.725 Nm^2, I_5E=5.67 Nm^2. As masses of the springs are less than 10% of the attached masses, they are omitted in calculation. In addition, the elastic supporting and damping of the system are also neglected (in translational model k_i=0,0001 Ns/m, avoiding divide by zero). Based on the 3D solid body model and using the relations (23)-(26), the stiffness coefficients $(s_i, i = 1 \dots 5)$ of the cantilever beams are computed for the translational modell

$$s_1 = 1/c_{11} \tag{27}$$

$$s_i = \frac{3I_iE}{b_i^3}, \quad (i = 2 \dots 5) \tag{28}$$

Using the Cramer's rule for inverse stiffness matrix and the relations (27) and (28), the amplitude of vibration A_{1m} of the cantilever beam with mass m_1 excited with frequency ω_g is obtained

$$A_{1m} = \frac{\begin{vmatrix} F_0 & 0 & 0 & 0 & 0 \\ 0 & C_{22}^{-1}-m_2\omega_g^2 & C_{23}^{-1} & C_{24}^{-1} & C_{25}^{-1} \\ 0 & C_{32}^{-1} & C_{33}^{-1}-m_3\omega_g^2 & C_{34}^{-1} & C_{35}^{-1} \\ 0 & C_{42}^{-1} & C_{43}^{-1} & C_{44}^{-1}-m_4\omega_g^2 & C_{45}^{-1} \\ 0 & C_{52}^{-1} & C_{53}^{-1} & C_{54}^{-1} & C_{55}^{-1}-m_5\omega_g^2 \end{vmatrix}}{\begin{vmatrix} C_{11}^{-1}-m_1\omega_g^2 & C_{12}^{-1} & C_{13}^{-1} & C_{14}^{-1} & C_{15}^{-1} \\ C_{21}^{-1} & C_{22}^{-1}-m_2\omega_g^2 & C_{23}^{-1} & C_{24}^{-1} & C_{25}^{-1} \\ C_{31}^{-1} & C_{32}^{-1} & C_{33}^{-1}-m_3\omega_g^2 & C_{34}^{-1} & C_{35}^{-1} \\ C_{41}^{-1} & C_{42}^{-1} & C_{43}^{-1} & C_{44}^{-1}-m_4\omega_g^2 & C_{45}^{-1} \\ C_{51}^{-1} & C_{52}^{-1} & C_{53}^{-1} & C_{54}^{-1} & C_{55}^{-1}-m_5\omega_g^2 \end{vmatrix}} \tag{29}$$

In Fig. 5 the amplitude diagram as the function of the excitation frequency is plotted.

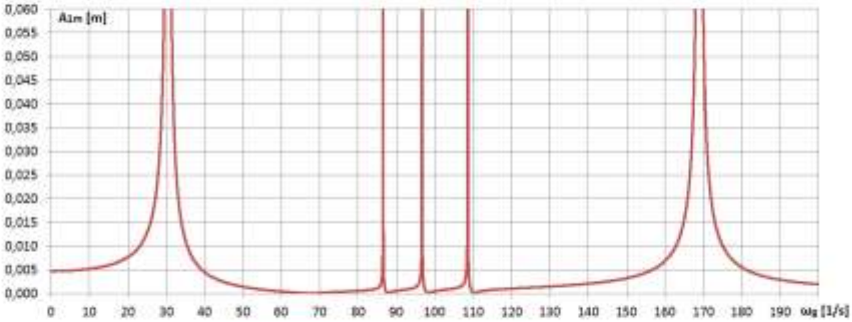


Figure 5

Amplitude – excitation frequency diagram for the 5-DoF cantilever-mass mechanism

According to the already published paper [18], the vibration amplitude A_{1t} of the mass m_1 in the 5-DoF 'wallpaper' model with translator motion (Fig. 1b) is

$$A_{1t} = \frac{F_0}{\sqrt{(s_1 - m_1\omega_g^2 - \sum_{i=2}^{5} m_i G_{i1}\omega_g^2 cos\varphi_i)^2 + (k_1\omega_g + \sum_{i=2}^{5} m_i G_{i1}\omega_g^2 sin\varphi_i)^2}} \tag{30}$$

where $G_{i1} = \sqrt{\frac{\left(2D_i\omega_g\omega_i^{-1}\right)^2 + 1}{\left(1 - \omega_g^2\omega_i^{-2}\right)^2 + \left(2D_i\omega_g\omega_i^{-1}\right)^2}}$, $\varphi_i =$

$arc\ tg \frac{\omega_g\omega_i^{-1}}{(2D_i)^{-1}\left(\omega_g\omega_i^{-1}\right)^{-2} - (2D_i)^{-1} + 2D_i}$,

$$D_i = \frac{k_i}{2m_i\omega_i}, \quad \omega_i = \sqrt{\frac{s_i}{m_i}} \tag{31}$$

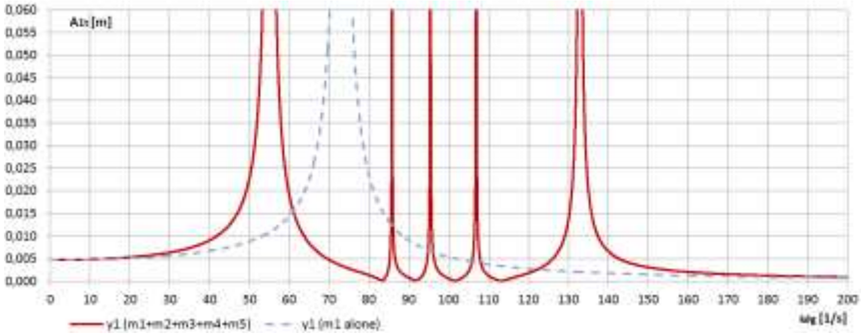


Figure 6

Amplitude - excitation frequency diagram for the 5-DoF translational model (solid line) and for 1-DoF translational model without absorber (dashed line)

In Fig. 6 the amplitude diagram as the function of the excitation frequency for the translator model (Fig. 1b) is plotted.

Comparing Fig. 5 and Fig. 6, it is visible that for both models (cantilever-mass and translational model, respectively) there are five resonances and four stopping positions (A_1=0) which are not at the same excitation frequencies. The three center resonances are almost at the same frequencies and the shape of the curves are similar. Positions of the first and the last resonances are different. The stopping frequencies of the cantilever model are smaller than of the other one.

In Fig. 7, the first six modal forms are presented for frequencies ω_1, ω_2, ... ω_6 of the 5-DoF cantilever-mass mechanism modelled as 3D solid body model (Fig. 2a). The FEA Mesh settings are: average element size is 0.100, minimum element size is 0.200, grading factor is 1.500, minimum turn angle is 60.00 deg, and curved mesh and elements are allowed. Local mesh control has to be in the interval 1.00 mm to the upper and lower side of the cantilever beam.

The results of the FEA can be seen as follows:

ω_1 – All the five masses vibrate in the same phase. This is the lowest resonance angular frequency of the main mass m_1.

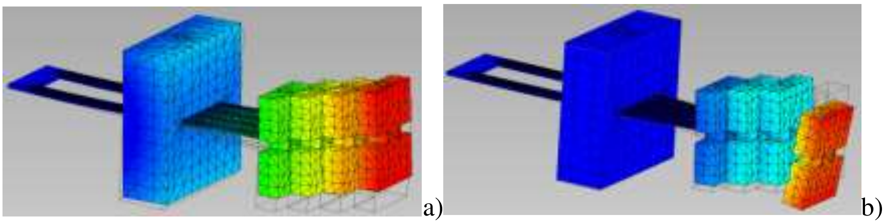
ω_2 – Mass m_1 stops, the mass m_5 attached with the longest stick resonates. This is the lowest angular frequency of vibration suppression.

ω_3 – Mass m_1 stops, the mass m_4 attached with the second longest stick resonates. This is the second lowest angular frequency of vibration suppression.

ω_4 – Mass m_1 stops, attached mass m_3 and mass m_2 resonate together. This is the third lowest angular frequency of vibration suppression. It is interesting that there are no two different resonance frequencies in the FEA where mass m_3 and mass m_2 resonate separately as it is expected.

ω_5 – System makes torsion movements. This is irrelevant for the present investigation.

ω_6 – The main mass and the attached masses vibrate in the opposite phase. This is the highest resonance angular frequency from the two models' point of view.



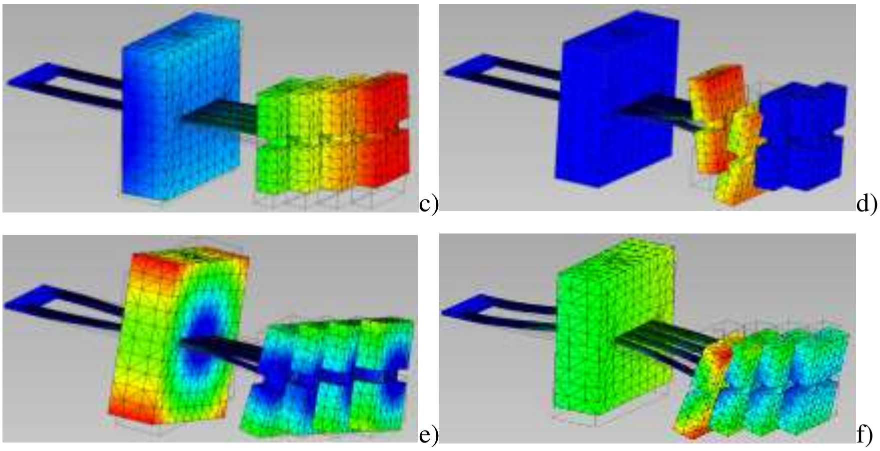


Figure 7

First six modal forms of FEA obtained by AutoDESK Inventor for frequencies: a) ω_1, b) ω_2, c) ω_3, d) ω_4, e) ω_5, f) ω_6

The angular frequencies ω_2, ω_3, ω_4 of Inventor FEA fit well with the first three stopping (A_1=0) angular frequencies of the translational model, while the first and sixth ones support the values for the cantilever-mass model (Table 1).

Table 1

Frequencies of vibration for translational and cantilever

	Translational model ω [rad/s]	Cantilever model ω [rad/s]	Inventor Finite Element Analysis for cantilever model ω [rad/s]		
A_1= 0 m	**83,21**	67,26	ω_2 →	**83,38**	m_5 moves
	91,65	**87,73**	ω_3 →	**92,05**	m_4 moves
	101,60	**98,24**	ω_4 →	**102,16**	m_2 and m_3 moves in opponent phase
	113,05	**110,28**	no matches		
			ω_5 →	135,40	Torsion
A_1 in resonance	55,06	**30,42**	ω_1 →	**29,34**	m_1 and the attached masses vibrates in phase
	85,76	**86,53**	no matches		
	95,39	**96,68**			
	106,88	**108,58**			
	132,87	**168,94**	ω_6 →	**159,72**	m_1 and the attached masses vibrates in opposite phase

According to the calculations, it is seen that in both models, the four attached masses stop the main mass at four different frequencies. However, some inconvenient resonances exist. Their decrease is obtained by using the dampers.

In Fig. 8, the amplitude-frequency diagram for various values of the damping in the translational model is considered. It shows the calculations of the translational model with relatively small damping values set for all (*k_i=0-5 Ns/m*), while the lowest critical damping value is much higher (*k_{5crit}=2 $m\omega_5$=83,21 Ns/m*).

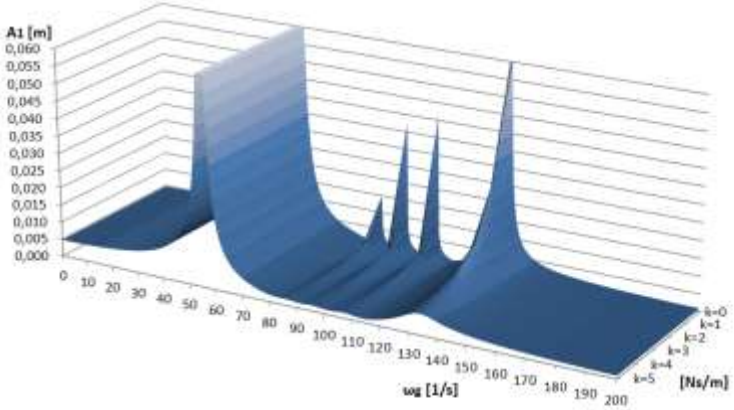


Figure 8
Amplitude-frequency-damping curve for m_1 at translational model

In Fig. 8, it is found that in the translational model (even for small damping) the amplitudes of vibration have the tendency of decrease in the excitation frequency range of 65 – 120 rad/s. Based on this result, it is expected that the same property is evident in cantilever-mass model.

Conclusions

In the paper, the multicantilever-mass mechanism for vibration suppression is considered. The physical model contains the system of clamped beams with attached masses which act as vibration absorbers. The mathematical model of the system is available for calculation of the amplitude of vibration for different excitation frequencies. Based on 1-DoF measurements, an elastic supporting theory was introduced for the calculating of the torque spring stiffness of clamping of the beam end. It is shown that the measured and calculated stiffness are in good agreement. In addition, it is concluded in the paper that a comparison between the cantilever mechanism suggested in the paper and the already existing translational model is possible. For certain frequency values, i.e. for the second, third and fourth frequencies, vibration is suppressed in both models. In addition, the amplitude-frequency plots for both models are similar in shape. Both models are in good agreement with numerical results matched with the Inventor FEA, too. However, it is obvious that the difference in results for the cantilever-mass system and translation mechanism is negligible.

Due to its simplicity (in comparison to the metastructure with translational motion of masses), the chosen cantilever system would be appropriate for measurements and future experimental investigation. The obtained experimental results for cantilever-mass mechanism should be treated for metastructure analysis.

As a final result of the investigation, the influence of damping in the 5-DoF metastructure was measured and calculated. It was highlighted that as the damping increases, the amplitudes of vibration for three medium resonance frequencies decrease dramatically.

For further investigation, the physical and mathematical model of the 'wallpaper' type metastructure's basic cell (with one coupled mass) can be developed.

Acknowledgement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Y. Cheng, J. Y. Xu, X. J. Liu, One-dimensional structured ultrasonic metamaterials with simultaneously negative dynamic density and modulus, Physical Review B, 2008, 77(4) 045134

[2] R. Zhu, X. N. Liu, G. K. Hu, C. T. Sun, G. L. Huang, A chiral elastic metamaterial beam for broadband vibration suppression, Journal of Sound and Vibration 333(10), 2014, pp. 2759-2773

[3] K. K Reichl, D. J. Inman Lumped mass model of a 1D metastructure for vibration suppression with an additional mass, Journal of Sound and Vibration 403, 2017, pp.75-89

[4] K. K. Reichl, D. J. Inman, Lumped mass model of a 1D metastructure with vibration absorbers with varying mass, in Sensors and Instrumentation, Aircraft/Aerospace and Energy Harvesting (eds. E.W. Sit), Vol. 8, Conf. Proceedings of the Society for Experimental Mechanics Series, 2019, pp. 49-56

[5] Y. R. Wang, T. Wen Liang, Application of lumped-mass vibration absorber on the vibration reduction of a nonlinear beam-spring-mass system with internal resonances, Journal of Sound and Vibration, 350,2015, pp. 140-170

[6] H. Sun, X. Du, P. F. Pai: Theory of metamaterial beams for broadband vibration absorption, Journal of Intelligent Material Systems and Structures, 21, July, 2010, pp. 1085-1101

[7] P. F. Pai, H. Peng, S. Jiang: Acoustic metamaterial beams based on multifrequency vibration absorbers, International Journal of Mechanical Sciences, 79, 2014, pp. 195-205

[8] P. F. Pai, Metamaterial-based broadband elastic wave undamped vibration absorber. Journal of Intelligent Material Systems and Structures 21(5), 2010, pp. 517-528

[9] L. Cveticanin, Gy. Mester, Theory of acoustic metamaterials and metamaterial beams: An overview, Acta Polytechnica Hungarica, 13(7), 2016, pp. 43-62

[10] Milton GW. (2007) New metamaterials with macroscopic behavior outside that of continuum elastodynamics. New Journal of Physics 9, 2007, 359:pp. 1-13

[11] H. Sun, X. Du, R. R. Pai, Theory of metamaterial beams for broadband vibration absorption, Journal of Intelligent Material Systems and Structures, 21 July 2010, pp. 1085-1101

[12] P. F. Pai, H. Peng, S. Jiang, Acoustic metamaterial beams based on multi-frequency vibration absorbers, International Journal of Mechanical Sciences, 79, 2014, pp. 195-205

[13] T. Wang, M. P. Sheng, Q. H. Qin, Multi-flexural band gaps in an Euler-Bernoulli beam with lateral local resonators, Physics Letters A, 380, 2016, pp. 525-529

[14] M. Askari et al.: Additive manufacturing of metamaterials: A review, Additive Manufacturing 36, 2020, 101562

[15] X. Yu et al., Mechanical metamaterials associated with stiffness, rigidity and compressibility: A brief review, Progress in Material Science 94, 2018, pp. 114-173

[16] H. Peng, P. F. Pai, Acoustic metamaterial plates for elastic wave absorption and structural vibration suppression, International Journal of Mechanical Sciences, 89, 2014, pp. 350-361

[17] H. Peng, P. F. Pai, H. Deng, Acoustic multi-stopband metamaterial plates design for broadband elastic wave absorption and vibration suppression, International Journal of Mechanical Sciences, 103, 2015, pp. 104-114

[18] P. Szuchy, 5-Degree-of-freedom systems in acoustic metamaterials, International Scientific Conference ETIKUM 2018,Proceedings, 6-8 December 2018, ISBN 978-86-6022-123-2 Novi Sad, Serbia, 2018

[19] J. D. Hobeck, C. M. V. Laurent, D. J. Inman, 3D printing of metastructures for passive broadband vibration suppression, 20th Int. Conf. on Composite Materials, Copenhagen, 19-24 July, 2015, pp. 1-8

[20] C. D. Pierce, C. L. Willey, V. W. Chen, J. O. Hardin, J. D. Berrigan, A. T. Juhl, K. H. Matlack, Adaptive elastic metastructures from magneto-active elastomers, Smart Material Structure 29, 2020, 065004, pp. 1-11

[21] B. C. Essink, D. J. Inman, Three-dimensional mechanical metamaterial for vibration suppression, Special Topics in Structural Dynamics & Experimental Techniques (Ed. N. Dervilia), Vol. 5, 2020, pp. 43-48

[22] L. Fan, Y. He, X. Chen, X. Zhao, Elastic metamaterial shaft witth a stack – like resonator for low-frequency vibration isolation, Journal of Physics D, Applied Physics, 53, 2020, 105101, pp. 1-9

[23] L. Cveticanin, M. Zukovic, Negative effective mass in acoustic metamaterial with nonlinear mass-in-mass subsystems, Communications in Nonlinear Science and Numerical Simulation, 51, 2017, pp. 89-104.

[24] L. Cveticanin, M. Zukovic, D. Cveticanin, On the elastic metamaterial with negative effective mass, Journal of Sound and Vibration, 2018, 436, pp. 295-309

[25] L. Cveticanin, M. Zukovic, D. Cveticanin, Influence of nonlinear subunits on the resonance frequency band gaps of acoustic metamaterial, Nonlinear Dynamics, 93(3), 2018, 1341-1354

[26] T. Wang, M. P. Sheng, Q. H. Qin, Multi-flexural band gaps in an Euler-Bernoulli beam with lateral local resonators, Physics Letters A, 2016, 380, pp. 525-529

Detecting Cyber Attacks with High-Frequency Features using Machine Learning Algorithms

Ahmet Nusret Özalp[1], Zafer Albayrak[2]

[1] Department of Computer Engineering, Karabuk University, Karabük, Turkey, ahmetnusretozalp@karabuk.edu.tr

[2] Department of Computer Engineering, Sakarya University of Applied Sciences, Sakarya, Turkey, zaferalbayrak@subu.edu.tr

Abstract: In computer networks, intrusion detection systems are used to detect cyber-attacks and anomalies. Feature selection is important for intrusion detection systems to scan the network quickly and accurately. On the other hand, analyzes performed using data with many attributes cause significant resource and time loss. In this study, unlike the literature studies, the frequency effects of the features in the data set are analyzed in detecting cyber-attacks on computer networks. Firstly, the frequencies of the features in the NSL-KDD data set were determined. Then, the effect of high-frequency features in detecting cyber-attacks has been examined with the widely used machine learning algorithms of Random Forest, J48, Naive Bayes, and Multi-Layer Perceptron. The performance of each algorithm is evaluated by considering Precision, False Positive Rate, Accuracy, and True Positive Rate statistics. Detection performances of different types of cyberattacks in the NSL-KDD dataset were analyzed with machine learning algorithms. Precision, Receiver Operator Characteristic, F1 score, recall, and accuracy statistics were chosen as success criteria of machine learning algorithms in attack detection. The results showed that features with high frequency are effective in detecting attacks.

Keywords: Attribute selection;Cyberattacks; Machine Learning; IDS;NSL-KDD; Anomaly detection

1 Introduction

The purpose of intrusion detection systems is to predict the attacks like infiltration, attack and malware in advance. From the security of information systems perspective, connection protocol bugs must be eliminated while the connection interfaces of network devices must be configured correctly. The detection of attacks is ensured by monitoring network as well as malicious traffic with port scans. For such a purpose, incoming and outgoing network packet information is watched over the network traffic. With the information received, data is collected to detect suspicious connections. With active and passive scans, vulnerabilities of IP address blocks, open ports, operating system information, running services, active devices and active hosts are discovered accordingly.

In general, three methods are used to detect cyber-attacks. These are signature-based attack detection, anomaly-based attack detection and hybrid-based attack detection systems. In the signature-based detection method, each attack is recorded by creating a dictionary (wordlist) with a uniquely defined signature. Each newly detected attack is stored in this dictionary. Thus, a defence system is formed upon known and discovered attacks [1]. The anomaly detection method evaluates whether there is an unusual situation or not by taking information packets from the traffic on the network. If an abnormal situation is detected, then the intrusion prevention system is activated. Anomaly detection-based systems can detect attacks that signature-based detection systems cannot detect. In order to increase detection success, hybrid systems have been developed by combining these two approaches. According to their usage areas, hybrid intrusion detection systems can be divided into two parts. The first part is an anomaly based hybrid intrusion detection system while the second one is a parallel-based intrusion detection system. Regarding intrusion detection systems, applications performed with machine learning (ML), data mining and deep learning (DL) algorithms are available in the literature. Besides, various data mining techniques are also used to detect abnormal conditions in network traffic [2-4]. In the classification of traffic, the focus has been on machine learning techniques [5-8]. Performance values such as accuracy, positive accuracy rate and detection time have been tested in intrusion detection applications using machine learning techniques [9-11]. In the detection of attacks, machine learning techniques provide higher accuracy over network traffic. The results showed that intrusion detection approaches using machine learning algorithms provide higher success compared to other methods [8, 12]. The effects of scanning methods on intrusion detection systems have been provided in Table 1. In addition, the access control lists of the packets sent and the packets returned through the firewall have been determined. The obtained information constitutes an important parameter for attacks so that checklists are created against port scanning attacks while the firewall is prompted to correctly detect port scanning operations. Devices that perform routing and filtering processes bypass certain source ports. Special rules are obtained for unwanted ports that prevent unauthorized access. According to these rules, data are collected on the network traffic. The status of the traffic on the network is monitored via the amount of collected data. If an abnormal situation is detected, then the determination of the attack method is requested. In network-based intrusion detection systems, fuzzy set theory [4, 13, 14], artificial neural networks [6, 10, 15], ML [14, 16, 17], and DL techniques are used to detect links that contain anomalies [18-19]. Log files and datasets for analysis are the main components for the studies to be conducted. While detecting anomalies with real-time packet analysis, it is difficult to determine the parameters such as performance and accuracy [2]. Datasets are used in cases such as excessive energy use and memory insufficiency in devices [7, 18]. They are also used as training and test data in studies where anomaly detection is performed with deep learning and ML algorithms. In such a case, the trained data precisely detect the attacks in real-time and provide information regarding the measures to be taken. All information added to the training data needs to be analyzed and controlled [10, 20].

Table 1

Network scanning attacks

Scanning Techniques	*Packet Sent*	*Port Open Close Detection*	*Returned Packet*	*Three-way Handshake*	*IDS Firewall Check*
TCP Connect/Full Open Scan	TCP	Yes	RST	Yes	Yes
Stealth Scan/Half-open Scan	TCP	Yes	RST	Yes	Yes
Inverse TCP Flag Scanning	FIN,URG, PSH	Yes	RST	Yes	Yes
Xmas Scan (Xmas Scanning)	IN,URG, PSH,TCP	Yes	Inverse TCP	Yes	Yes
ACK Flag Probe Scanning	TCP /ACK	Yes	RST	Yes	Yes
IDLE/IPID Header Scan	TCP,SYN	Yes	RST/SYN ACK/ RST	Yes	Yes

Thus, learning is provided with the tagged data. Increases in the number of users and devices, as well as difficulties in detecting real-time attacks cause hardware inadequacies and cause higher costs in detecting attacks by devices. In this study, more effective detection of cyber-attacks by attribute selection is proposed as it contributes to more effective cyber-attacks detection with high-frequency feature selection in datasets carrying attack information.

In this study, the main contribution of our study to the literature is the analysis of both anomaly-based attacks in the network and DDOS, U2R, R2L attacks with high classification success machine learning algorithms. Unlike previous studies, high-frequency features were determined as a result of the sequencing, and the detection rates of the attacks were analyzed by machine learning algorithms.

• 41 features with dataset were first dimensioned by using One-R, Chi-square (Chi-S), Correlation-Based Self-Attribute Selection (CBS), Symmetrical Uncertainty Coefficient (SUC), Gain Rate (GR), Information Gain (IG) selection methods. Unlike the studies in the literature, the frequencies of the features to be used in classification were determined. Then, the effects of high-frequency features in detecting different attacks were examined. In particular, 4, 5, 6, 29 and 30 valued attributes were effective in detecting anomalies, while 3, 4, 5 and 6 valued attributes were found to be effective in detecting DoS attacks.

• Feature vectors were classified by using Random Forest, J48, Naive Bayes, and Multi-Layer Perceptron algorithms. For the classification, accuracy, time, positive correct rate, and positive false rate were considered for the performance criteria of the algorithms. Besides, the effect of the attributes was provided while calculating the performance criteria of five different attacks.

• The performances of machine learning algorithms were compared according to the criteria of Precision (P), False Positive Rate (FPR), True Positive Rate (TPR), Accuracy (Acc) according to the high-frequency attributes.

The content of the article is organized as follows. In Section 2, information about the studies on the subject is provided. Feature selection methods are explained in Section 3. In Section 4, information about the machine learning model approaches used in intrusion detection is expressed as well as the classification algorithms used in the study. In Section 4, the results of the analysis are also evaluated while examining the effects of parameters on anomaly detection in computer networks. Finally, the results are summarized in Section 5.

2 Background

Most of the approaches such as fuzzy logic and data mining in detecting cyber-attacks have not yielded the desired results in larger datasets. As there are many attributes in the data collected from the computer network, the classification and detection of attacks cause a waste of time [3]. On the other hand, the information contained in the attributes is important for the accuracy of the classification. If the number of attributes is low, then the classification quality decreases while the error rate in the detection of attacks increases due to the generalizations. Besides, data processing time increases and real-time attacks become difficult to detect if the number of attributes is high. In attack detection systems, attribute sizing operations in the dataset are observed to decrease the attack detection time and increase the accuracy [4, 5]. Some studies perform machine learning and deep learning approaches in attack detection systems. The most recent ones are listed in Table 2 for to features of attribute selection and dataset usage as well as machine learning and deep learning algorithms. The current study is also compared with the existing studies according to the same criteria. Anomaly detection studies come to the forefront in studies using data-based techniques in detecting attacks [21, 22]. In machine learning algorithms, the NSL-KDD dataset is preferred due to its high number of attributes and its reliability in attack detection scenarios [17, 21]. Attribute numbers and selection methods are important parameters in detecting anomalies in machine learning algorithms [23, 24]. As a result of the classification according to the number of attributes selected, anomaly detection percentages between 97% and 99% were obtained [25, 26]. According to the analysis of 41 attributes in the proposed DNN approaches, the trained data was determined to be insufficient due to the increase in the number of classes [5]. The above-mentioned features reveal the difficulties of collecting packets from network traffic in real-time attack detection as well as their performance reduction effect [6-20, 27-30]. In this study, the attribute selection performance was examined over the NSL-KDD dataset as the first step. Machine learning algorithms were preferred for classification algorithms after considering fuzzy logic, data mining, machine learning, and deep learning algorithms. Machine learning algorithms tend to represent high performance according to criteria such as accuracy, precision, and time in classification for the attributes selected on high-dimensional datasets [16, 20, 31, 32]. As the size of the data increase, the difficulties in interpreting the data reveal new approaches such as deep learning [23, 28, 33, 34]. In machine learning, it is desirable to store, change, and process the data in a suitable format to make it meaningful. The data are converted to matrix format and processed in tables before the estimation is performed. When deep learning approaches are examined, an appropriate model is designed [20, 29, 31]. By forming a model with the existing parameters, the suitability of the available data for the model is examined. Deep learning provides successful results in areas such as natural language processing, anomaly detection, and pattern recognition. In order to apply deep learning, the problem must be defined correctly as the first step. The mathematical model of the problem can be created while the improvement in the system can thus be observed by applying the relevant techniques. The dataset of this study has 24 different network attack examples in 4 categories. DoS (denial of service) is the name given to attacks to prevent network access. Probe (search) is defined as the scanning of IP and ports to detect vulnerabilities in the target. In R2L (remote to local), the attackers do not have the privilege to log in but can send the packet to the destination.

Table 2
Comparison with related research work

Study	*Attribute Selection*	*Machine Learning Approach*	*Deep Learning Approach*
Xin et al. [6]	No	Yes	Yes
Da Costa et al. [7]	Yes	No	No
C.bouni et al. [8]	Yes	Yes	Yes
Berman et al. [9]	Yes	Yes	Yes
Mazini M. et al. [10]	No	Yes	Yes
Sultana et al. [11]	No	Yes	No
Ferrag et al. [12]	No	Yes	Yes

On the U2R (user to root), they are able to monitor password entries to gain aggressive access. Even access right in standard user mode is possible, authorized users try to access. The 41 attributes in the dataset can be evaluated individually as well as under 4 categories according to the attack types.

Table 3
Comparison of attribute selection with relevant research studies

Author(s)	***Datasets***	***Approaches***	***Attribute selection type***
S.Thaseen et al. [14]	NSL-KDD	Weighted majority voting	Chi-S
Kasongo et al. [15]	NSL-KDD, UNSW-NB15	Two-stage ensemble	CBS
Mazini et al. [10]	NSL-KDD, ISCX 2012	Ada boost, Naïve Bayes	Bee colony
Verma et al. [16]	Private	Boosted tree,NB	-
Pham et al. [17]	NSL-KDD	Bagging,J48	GR
Aljawarneh et al. [18]	NSL-KDD	Majority voting, MLP	IG
Zaman et al. [19]	Kyoto 2006+	Majority voting	Information entropy
Al-Jarrah et al. [20]	NSL-KDD, Kyoto+	Random forest, Naïve Bayes	-
Vigneswaran et al. [21]	KDD Cup99	Random forest	-

These are the attacks made according to Transmission Control Protocol (TCP) connection characteristics, time-tagged attacks of two seconds, attacks lasting more than two seconds, and attack attributes based on content information [15]. The dataset is used in many academic studies, especially because of its high potential for anomaly detection and detection of new attacks. In literature studies have been conducted with this dataset using algorithms such as Support Vector Machine (SVM), K-Means, Random Forest, and J48 [15, 37]. Table 3 lists the studies conducted according to the methods used in attribute selection. Considering other methods, information gain was also preferred in this study. The reason behind this preference is the decision making capability in more diverse datasets although it does not have much data in the information acquisition method [11, 35, 40-43]. In the case of larger datasets, the need of being supported by other selection methods is required [21, 26, 35, 38, 39, 44-45].

3 Attribute Selection

For the attribute selection procedure, 10 attributes for NSL-KDD are selected according to the sorting criteria. Then, the selected attributes are classified accordingly and transferred to the model. The model suggested in this study is presented in Figure 1. A total number of 10 attributes were selected for NSL-KDD by reducing the received datasets into subsets and sequencing them with attribute selection methods. As presented in Figure 1, the size of the training set should be redefined and brought into the appropriate evaluation range. In this process, the rate of gain, correlation-based attribute selection, information gain, chi-square, symmetric uncertainty coefficient, and One-R are selected as the attribute determination method. Since the best attributes are determined at this stage, the importance of this stage is inevitable. In particular, reducing the size of the collected data for anomaly detection reduces the burden during attack detection.

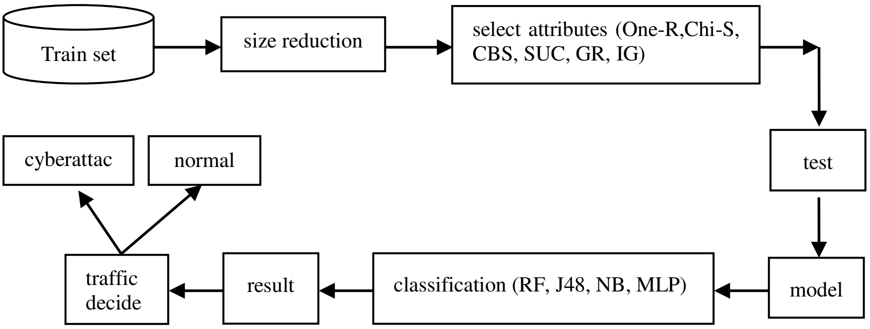

Train set
size reduction
select attributes (One-R,Chi-S, CBS, SUC, GR, IG)
test
model
classification (RF, J48, NB, MLP)
result
traffic decide
cyberattac
normal

Figure 1
The model structure

The success of the system increases with the attributes obtained from known attack types. The received dataset is divided into subsets by methods such as random and complete search. Thus, a sub-dataset for evaluation is formed. Depending on the selection, dependent or independent criteria are determined. During this stage, the process continues until enough subsets are formed to determine the best subset. Since there is uncertainty, entropy is used during attribute selection. The following criteria ought to be considered for the selection of the attributes.

3.1 Correlation-Based Self-Attribute Selection (CBS)

CBS is based on determining clusters that are not directly related to each other since it has a filtering logic. The low correlated attributes are eliminated and the data with high frequencies are used by the following formula [19].

$$M_s = \frac{rfc}{\sqrt{l + k(k-1)rff}} \tag{1}$$

M_s= Merit value of the subset S with k features

rcf= Correlation between the class tag and the associated attribute

rff= Correlation of attributes.

3.2 Chi-square (CS)

CS is a statistical method where initial values observed by classes are calculated based on X2 statistics. In the next step, a selection is performed according to the number of attributes minus 1 in the dataset depending on its importance status. If the expected frequency value matches, x2 approaches zero while it indicates incompatibility otherwise according to the following formula [20].

$$x^2 = \sum_{i=1}^{n} \frac{(o-e)^2}{e} \tag{2}$$

n: number of attributes in the dataset

o: Observed frequency value for the ith attribute

e; Expected frequency value for the ith attribute.

3.3 One-R (1-R)

Each attribute in the dataset allocated for training is classified according to the determined rule. Depending on the error rate, sorting is performed according to the most frequently encountered attributes. Defines a rule for the entire prediction, each value of the prediction made, and the frequency of each value in the class is counted. The class with the highest frequency is determined while the corresponding prediction is added to the rule. The total error is calculated and the one with the lowest error rate is selected [23].

3.4 Symmetrical Uncertainty Coefficient (SUC)

In order to eliminate and eventually normalize the negative cases arising in the information gain method, the entropies of the attributes sampled as X and Y are added where the SUC is defined as [24].

$$\text{Coefficient} = \left(2 \frac{\text{IG}}{\text{H(y)} + \text{H(x)}} \right) \tag{3}$$

3.5 Information Gain (IG)

Information gain is a way of normalizing the negative parts in symmetrical uncertainty gain and is based on entropy. The X property and the Y property vary depending on their respective values (Eq. 4). The biggest drawback of this method is that it may make decisions in favor of more diverse datasets although it does not have much data [25].

$$\mathrm{IG} = \mathrm{H(y)} - \mathrm{H(y}\|\ \mathrm{x)} \tag{4}$$

By measuring the information gain according to the class, the property value is examined.

$$\mathrm{IG(class, attribute)} = \mathrm{H(class)} - \mathrm{H(class|\,attribute)} \tag{5}$$

3.6 Gain Rate (GR)

Gain rate is used to normalize the information gained method to minimize the resulting diversity by [26].

$$\mathrm{G}R = \left(\frac{\mathrm{IG}}{H(x)}\right) \tag{6}$$

4 Experimental Work

In this study, 6 different attribute determination methods of the NSL-KDD dataset are used. 20% of the data set features were selected. This ratio was also taken into account during the determination of the training and test dataset. In the studies conducted in the literature, no specific reason for the number of selected features has been revealed [25, 44]. In this study, the 10 most successful attributes were selected according to their performance order for each method. Table 4 lists the attributes chosen as the basis for ordering. When the studies conducted with the NSL-KDD dataset are examined [32], it was observed that the number of selected attributes, attribute selection method, and classification approaches are different. 10 attribute names, their numbers in the dataset, and their frequencies obtained as a result of the attribute selection methods are shown in Table 5. The list in the table indicates that the frequencies of the attributes numbered 4, 5, 6, 12, 29, and 30 are high. At the end of the feature selection in Table 5, features with high frequency are observed. As a result of the feature selection procedure for 6 different methods, the order of each attribute was determined. The features with the highest frequency in this ranking are provided in Table 6. Here, flag data checks the connection status while src-byte and dst-byte check the link status of source and destination points during the connection. Diff-srv-rate and same-srv-rate attributes represent the connection status of the attacker to the same point. These parameters were obtained to be effective in detecting 5 different attack types in the dataset as well as detecting anomalies in the network due to their high frequency. These attribute subsets were analyzed using 4 different classification algorithms such as Naive Bayes, J48, Multi-Layer Perceptron, and Random Forest respectively [46]. No specific intrusion detection reference measure is obtained as it is decided by the classification and modeling of current attack types. In this study, precision, false-positive rate, accuracy, and true-positive rate were considered the criteria for detecting the attacks. The common feature of the high-frequency attributes is their usage possibility in the detection of DoS attacks and anomaly status in the network. The similarity rate of the features used in the detection of the other three types of attacks is obtained to be 73%.

Table 4

NSL-KDD Dataset results obtained with attribute selection methods and their ranking

One-R			**Correlation-Based Self-Attribute Selection**		
Attributes			**Attributes**		
Ranked	**Number**	**Name**	**Ranked**	**Number**	**Name**
96.374	5	src_bytes	0.747	3	service
91.558	3	service	.0725	5	src_bytes
90.900	6	dst_bytes	0.695	12	logged_in
88.090	4	flag	0.692	4	flag
87.380	30	diff_srv_rate	0.691	6	dst_bytes
87.324	29	same_srv_rate	0.634	29	same_srv_rate
85.426	34	dst_host_s_sr_rate	0.595	30	diff_srv_rate
85.010	33	dst_host_srv_count	0.576	25	serror_rate
83.920	35	dst_hst_di_srv_rate	0.563	26	srv_serror_rate
82.947	12	logged_in	0.531	33	dst_host_srv_count
Gain Ratio Feature			**Chi-Square**		
Attributes			**Attributes**		
Ranked	Number	Name	Ranked	Number	Name
0.418	12	logged_in	109.922	5	src_bytes
0.373	26	srv_serror_rate	93.032	3	service
0.339	4	flag	87.820	6	dst_bytes
0.332	25	serror_rate	75.735	4	flag
0.332	39	dst_host_srv_s_rate	74.897	30	diff_srv_rate
0.267	30	diff_srv_rate	73.850	29	same_srv_rate
0.264	38	dst_host_serror_rate	69.215	33	dst_host_srv_count
0.258	6	dst_bytes	67.900	34	dst_host_s_sr_rate
0.231	5	src_bytes	62.343	35	dst_hst_di_srv_rate
0.224	29	same_srv_rate	60.430	12	logged_in

Symmetrical Uncertainty Coefficient			**Information Gain Attribute**		
Attributes			**Attributes**		
Ranked	**Number**	**Name**	Ranked	Number	Name
0.411	12	logged_in	0.816	5	src_bytes
0.411	4	flag	0.671	3	service
0.377	6	dst_bytes	0.633	6	dst_bytes
0.367	26	srv_serror_rate	0.519	4	flag
0.362	39	dst_host_srv_s_rate	0.518	30	diff_srv_rate
0.360	25	serror_rate	0.509	29	same_srv_rate
0.360	5	src_bytes	0.475	33	dst_host_srv_count
0.353	30	diff_srv_rate	0.438	34	dst_host_s_sr_rate
0.320	38	dst_host_serror_rate	0.410	35	dst_hst_di_srv_rate

Table 5
Attribute frequencies

Attribute Number	*Attribute Name*	*Frequency*
3	service	4
4	flag	6
5	src-bytes	6
6	dst-bytes	6
12	logged-in	5
25	serror-rate	3
26	srv-serror-rate	3
29	same-srv-rate	6
30	diff-srv-rate	6
33	dst-host-srv-count	3
34	dst-host-same-srv-rate	3
35	dst-host-diff-srv-rate	3
38	dst-host-serror-rate	3
39	dst-host-srv-serror-rate	2

Table 6
High-frequency features

No	*Attribute name*	*Description*	*Sample Data*
4	Flag	Connection status Normal or Error	SF
5	src-bytes	Number of data bytes transferred from source to destination in a single connection	491
6	dst-bytes	Number of data bytes transferred from destination to source in a single connection	1
29	same srv-rate	Percentage of connections to the same service among the connections aggregated	1
30	diff-srv-rate	Percentage of connections to different services among the connections collected	1

4.1 Performance Criteria

The accuracy of the classification process is measured by the "Confusion Matrix". This matrix provides an understanding of the probability outcomes in classification. If there is a dual classification such as anomaly detection, the labelling is done as normal and abnormal. Four conditions arise in binary guessing. True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). While measuring the accuracy of the model put forward accordingly, True Positive Rate (TPR) and False Positive Rate (FPR) are used.

Precision (P): Precision refers to the possibility of making an accurate estimate with the data obtained and defined as [33],

$$\mathrm{P} = \left(\frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \right) \tag{7}$$

False Positive Rate (FPR): It is the rate of classifying the obtained data with an erroneous approach formulated as [34],

$$\mathrm{FPR} = \left(\frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} \right) \tag{8}$$

True Positive Rate (TPR): It is the number of correct samples included in the positively grouped class defined as, [36].

$$\mathrm{TPR} = \left(\frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \right) \tag{9}$$

Accuracy (Acc): Of the total sample in the dataset, it is the percentage of the correctly estimated sample formulated as [35],

$$\mathrm{Acc} = \left(\frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + TN + FP + \mathrm{FN}} \right) \tag{10}$$

Receiver Operator Characteristic (ROC): ROC is used to calculate cost sensitivity in classification processes. It is obtained by drawing the curve between False Positive Rate and True Positive Rate in the detection of anomalies. In this way, the performance of the algorithm used as a classifier is compared between error costs and class distributions. The area under the curve shows the accuracy of the model estimation to be obtained as a result of the classification [33].

F-1 Score: It is an accuracy parameter for the test. It is calculated according to the sensitivity (P) and recall [37].

$$\mathrm{F1\text{ - }Score} = \left(\frac{2\mathrm{TP}}{2\mathrm{TP} + FP + \mathrm{FN}} \right) \tag{11}$$

Recall: It is the ratio of positive correct predictions to samples in the positive grade [38].

$$\text{Recall} = \left(\frac{\text{TP}}{\text{TP} + \text{FP}} \right) \tag{12}$$

4.2 Machine Learning Approaches with Intrusion Detection Systems

While optimization provides a way to minimize loss functionality for deep learning and machine learning, the goal of both methods is fundamentally different. The former is mainly concerned with minimizing a goal while the latter is concerned with finding a suitable model when a limited amount of data is available. The purpose function of the optimization algorithm is to reduce the training error. When it is a loss function, it is usually based on the training dataset. However, the purpose of statistical inference is to reduce the generalization error. This indicates that machine learning and deep learning approaches can be used in intrusion detection systems. Structurally, intrusion detection systems have to respond quickly to cyber-attacks. The use of machine learning algorithms in intrusion detection systems can be evaluated by using classification method, scaling method, and both classification and scaling methods. The attributes selected from the dataset are processed by machine learning algorithms during the classification stage. The techniques used in the classification stage provide the appropriate model creation. The studies about deep learning indicate that it needs improvement although it provides successful results especially in detecting anomalies [23]. Machine learning approaches on the other hand provide successful results in detection and prevention.

4.2.1 Random Forest Algoritm (RF)

Decision trees form a tree structure for classification models. Information gain and entropy metrics are important parameters in this algorithm. First, a decision tree is created with the learning set. In the next step, each new input data is determined as a class label. The Random Forest algorithm, which is also known as the decision tree classification, is the classification algorithm used to detect cyber-attacks. It is frequently used in anomaly detection, analysis of malware and vulnerability analysis in the detection of cyber attacks [26]. New branches are created by comparing each node that makes up the tree and the attributes divided into subsets while the leaves of the tree are expressed as a class. The biggest advantage of this algorithm over other algorithms is the presence of fewer parameters. It does not take unnecessary action against the abnormal data in the dataset so that it works with lower loads [28]. The classifier is defined as,

$$h(x,k), k = 1,2,...i \tag{13}$$

h: Classifier θk: random vector x: tree class tag

4.2.2 Naive Bayes Algorithm (NB)

The Naive Bayes algorithm is a Bayes' approach for classification where each attribute pair is processed independently. It evaluates the data independently. The aim is to have an equal effect on the result for each parameter. Structurally, it is a very simple and fast algorithm. It is one of the preferred methods for detecting cyber-attacks [26]. It can also provide results in a short time since it requires less training data for sampling points. NB reduces the problem of separator classes to find classes with conditional marginal densities. For this reason, representing the probability that a given sample is one of the possible target classes. Unless it contains inputs associated with each other, NB performs well against other algorithms.

$$P(c \mid x) = \frac{P(x \mid c) * P(c)}{P(x)} \tag{14}$$

P(c | X)=P(x_1 | c)*P(x_2 | c)*…*P(x_n | c)*P(c)

P(c):Class Prior Probability P(c | x):Posterior Probability

P(x):Predictor Prior Probability P(x | c):Likelihood

4.2.3 Multi Layer Perceptron (MLP)

Multi-Layer Perceptron is a feed-forward neural network, which consists of at least three layers as input, hidden, and output layers [27]. Except for input nodes, every node uses a non-linear activation function. It uses a feedback supervised learning technique for training. In this respect, it can be used in a non-linear system to distinguish the desired data.

$$f(x) = \sum_{i=1}^{m} (wi * xi) + b \tag{15}$$

m: The number of neurons in the previous layer w: random weight

x: input value b: random bias

4.2.4 J48 Algorithm

J48 is an algorithm developed by Ross Quinlan and considered the continuation of the ID3 algorithm. As it can create a decision tree, it is used as a statistical classifier in the structural sense [28]. In this classifier, a flowchart in the form of a tree model is created while the problem is tried to be solved based on prediction.

The nodes in the tree indicate the samples taken for entry while the leaves represent the estimates based on this entry.

5 Experiment and Results

After selecting the attributes of the dataset in detecting the attacks, the classification process was performed. The accuracy rate of classification according to attack types is expected to be high. In particular, the correctness of the classification in the detection of anomalies ensures correct interpretation of the features in the dataset. The algorithms listed in Table 7 were chosen due to their high classification rate. Successful results were obtained in P, FPR, Acc, and TPR percentages depending on the time in the anomaly detection with the selected features. Table 7 represents the anomaly classification algorithm performances conducted on the NSL-KDD dataset. NSL-KDD data set with 58630 anomalies and 67343 normal traffic data was used during training while 12833 anomaly data and 9711 normal data were used as test data. Random Forest, J48, Naive Bayes and MLP-CNN showed success rates of 99.76%, 98.45%, 93.34% and 91.34%, respectively, in the classifications made with the selected 10 features. The success in this classification rate was used to determine the attributes used for the detection of attacks. Classification results were evaluated according to P, FPR, Acc and TPR, by examining the criteria specified in the literature. Machine learning methods were compared using false alarm rate, accuracy and detection rate to detect anomalies in the network. It has been observed that the high rate of classification success also affects the success of the algorithm in detecting anomalies.

Table 7
Anomaly detection rates with J48, MLP, RF, and NB classification approaches concerning the obtained feature selections.

		J48 Algorithm					**Multi-Layer Perceptron (MLP)**				
		Correctly Classified Instances: 99.7817 %					**Correctly Classified Instances: 98.4354 %**				
Attribute & Methods		**Time (sec)**	**P %**	**FPR %**	**Acc %**	TPR %	**Time (sec)**	**P %**	**FPR %**	**Acc %**	**TPR %**
5,3,6,4,30,29,34,33, 35,12	**One-R**	3.34	76.86	4.23	93.56	93.52	19.65	69.29	4.17	90.86	83.50
12,26,4,25,39,30,38,6,5,29	**GR**	4.11	73.20	6.58	90.23	89.42	30.43	77.41	7.12	94.52	91.47
5,3,6,4,30,29,33,34, 35,12	**CS**	3.01	74.52	4.23	93.46	91.12	18.40	68.86	5.21	90.14	83.20
5,3,6,4,30,29,3	**IG**	3.05	74.56	4.74	93.20	90.36	19.97	68.27	5.74	91.01	82.98

3,34, 35,38											
12,4,26,6,39,25,5,30, 38,29	**SUC**	3.98	76.59	7.52	93.76	89.93	23.43	65.98	7.51	99.52	81.26
3,5,12,4,6,29,30,25, 26,33	**CBS**	5.43	74.86	7.41	95.47	88.63	22.50	75.58	6.27	90.74	80.37
All Attributes	**-**	6.78	73.47	4.69	90.45	98.45	45.86	72.41	4.12	91.38	91.34
		Random Forest (RF)					**Naive Bayes (NB)**				
		Correctly Classified Instances: 99.9174 %					**Correctly Classified Instances: 90.4178 %**				
Attribute & Methods		**Time (sec)**	**P %**	**FPR %**	**Acc %**	TPR %	**Time (sec)**	**P %**	**FPR %**	**Acc %**	**TPR %**
5,3,6,4,30,29, 34,33,35,12	**One-R**	2.90	78.65	3.23	94.95	95.43	5.20	75.38	5.89	91.58	89.36
12,26,4,5,39, 30,38,6,5,29	**GR**	4.11	74.82	5.43	91.23	91.23	4.78	71.24	8.23	90.56	88.26
5,3,6,4,30,29, 33,34,35,12	**CS**	3.01	76.78	3.21	94.65	93.54	3.97	72.28	5.27	90.26	89.78
5,3,6,4,30, 29,33,34,35,38	**IG**	3.05	76.71	3.28	94.89	93.87	3.98	72.56	5.23	90.29	89.87
12,4,26,6,39, 25,5, 30,38,29	**SUC**	3.98	78.42	6.40	95.54	91.20	4.21	74.36	8.54	90.56	88.25
3,5,12,4,6,29, 30,25,26,33	**CBS**	5.43	76.43	6.54	96.54	89.65	6.23	72.57	9.23	91.14	87.41
All attributes	**-**	6,78	75,43	3,45	95,45	99,76	8,30	70,29	4,69	91,45	93,34

Table 8

Performance of probe attack, U2R, R2L, and DoS attack types in machine learning classifiers

	Algorithms	*P*	*ROC*	*F1-Score*	*Re-call*	*Acc (%)*
Probe Attack	Multi-Layer Perceptron (MLP)	0.954	0.996	0.998	0.998	98.510
	Naive Bayes	0.986	0.976	0.961	0.971	90.398
	Random Forest	**0.999**	**1.000**	**1.000**	**1.000**	**99.952**
	J48	0.994	0.999	0.999	1.000	99.951
	Algorithms	*P*	*ROC*	*F1-Score*	*Re-call*	*Acc (%)*
User Root Attack	Multi-Layer Perceptron (MLP)	0.995	0.995	0.995	0.995	99.210
	Naive Bayes	0.999	0.949	0.961	0.943	88.859
	Random Forest	0.999	**0.998**	0.997	**0.998**	**99.859**
	J48	**1.000**	0.937	**0.998**	**0.998**	99.674
	Algorithms	*P*	*ROC*	*F1-Score*	*Re-call*	*Acc (%)*
Remote to Local Attack	Multi-Layer Perceptron (MLP)	0.997	0.996	0.992	0.992	99.814
	Naive Bayes	0.999	0.957	0.935	0.889	98.928

	Random Forest	**0.999**	**0.999**	**0.999**	**1.000**	**99.999**
	J48	0.998	0.995	0.998	0.999	99.997
	Algorithms	***P***	***ROC***	***F1-Score***	***Re-call***	***Acc (%)***
DoS Attack	Multi-Layer Perceptron (MLP)	0.954	0.841	0.948	0.998	95.752
	Naive Bayes	0.979	0.909	0.951	0.914	94.178
	Random Forest	**1.000**	**0.999**	**0.999**	**0.999**	**99.842**
	J48	0.995	0.667	**0.999**	**0.999**	99.774

From the results of experiments, it is seen that the number of features selected in the NSL-KDD dataset and the classification algorithm attacks affect the detection rate. Performance varies depending on the dataset size and the number of attributes selected. In previous studies, feature selection and the number of features were taken into consideration rather than high-frequency features. In previous studies, feature selection and the number of features were taken into consideration rather than high-frequency features. This situation was seen to directly affect the classification percentages of machine learning algorithms. With the attack detection study conducted with high-frequency features, the Random Forest algorithm was 1.7%; 0.97% of the J48 algorithm; 0.86% better results of NB algorithm, and 1.3% better results of MLP algorithm were obtained.

Conclusion

The results obtained in this study indicated that Random Forest Algorithm provides high performance in terms of classification and accuracy in the case of high-frequency features. Random Forest is followed by J48, NB, and MLP respectively. The most important feature identification function among datasets, which is one of the advantages of the Random Forest algorithm, has increased its success in attack analysis with the selection of high-frequency features. It has been observed that the success of the MLP algorithm used in linear functions in detecting cyberattacks is lower than other algorithms. When the features with high frequency are analyzed with machine learning algorithms, it is observed that especially the Random Forest algorithm produces 1.7% more accurate results.

References

[1] Meng, Weizhi, Wenjuan Li, and Lam-For Kwok, EFM: enhancing the performance of signature-based network intrusion detection systems using enhanced filter mechanism, Computers & Security 43 (2014), pp. 189-204

[2] Alabadi, Montdher, and Zafer Albayrak, Q-Learning for Securing Cyber-Physical Systems: A survey, In 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), IEEE, 2020, pp. 1-13

[3] Nazir, Anjum, and Rizwan Ahmed Khan., Network Intrusion Detection: Taxonomy and Machine Learning Applications, Machine Intelligence and

Big Data Analytics for Cybersecurity Applications. Springer, Cham, 2021, pp. 3-28

[4] Dwivedi, Shubhra, Manu Vardhan, and Sarsij Tripathi., Distributed denial-of-service prediction on iot framework by learning techniques, Open Computer Science 10.1 (2020) pp. 220-230

[5] Alabadi, M., & Celik, Y. (2020, June) Anomaly detection for cyber-security based on convolution neural network: A survey. In 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA) (pp. 1-14) IEEE

[6] Dua, Mohit, Attribute selection and ensemble classifier based novel approach to intrusion detection system, Procedia Computer Science 167 (2020) pp. 2191-2199

[7] Amiri, Fatemeh, et al, Mutual information-based feature selection for intrusion detection systems, Journal of Network and Computer Applications 34.4 (2011) pp. 1184-1199

[8] Akhter, A. F. M., Ahmed, M., Shah, A. F. M., Anwar, A., Kayes, A. S. M., & Zengin, A. (2021). A blockchain-based authentication protocol for cooperative vehicular ad hoc network. Sensors, 21(4) 1273

[9] Xin, Yang, Lingshuang Kong, Zhi Liu, Yuling Chen, Yanmiao Li, Hongliang Zhu, Mingcheng GAO, Haixia Hou, and Chunhua Wang, Machine learning and deep learning methods for cybersecurity, IEEE Access 6 (2018) pp. 35365-35381

[10] Da Costa, Kelton AP, et al, Internet of Things: A survey on machine learning-based intrusion detection approaches, Computer Networks 151 (2019) pp. 147-157

[11] Chaabouni, Nadia, et al., Network intrusion detection for IoT security based on learning techniques, IEEE Communications Surveys & Tutorials 21.3 (2019) pp. 2671-2701

[12] Berman, Daniel S., et al, A survey of deep learning methods for cyber security, Information 10.4 (2019) p. 122

[13] Dey, Samrat Kumar, and Md Rahman, Effects of machine learning approach in flow-based anomaly detection on software-defined networking, Symmetry 12.1 (2020) p. 7

[14] Otor, Samera Uga, et al., An improved bio-inspired based intrusion detection model for a cyberspace, Cogent Engineering 8.1 (2021) p. 1859667

[15] Alghamdi, Mohammed I., Survey on Applications of Deep Learning and Machine Learning Techniques for Cyber Security, International Journal of Interactive Mobile Technologies 14.16 (2020)

[16] Sultana, Nasrin, et al., Survey on SDN based network intrusion detection system using machine learning approaches, Peer-to-Peer Networking and Applications 12.2 (2019) pp. 493-501

[17] Ahmed, M., Moustafa, N., Suaib Ahther, A. F. M., Rezzak, I., Surid, E., Anwar, E., Shanen Shah, A.F.M & Zengin, A. (2021) A Blockchain-Based Emergency Message Transmission Protocol for Cooperative VANET. IEEE Transactions on Intelligent Transportation Systems (pp. 1-10)

[18] Revathi, S., and A. Malathi, A detailed analysis on NSL-KDD dataset using various machine learning techniques for intrusion detection, International Journal of Engineering Research & Technology (IJERT) 2.12 (2013) pp. 1848-1853

[19] Mazini, Mehrnaz, Babak Shirazi, and Iraj Mahdavi, Anomaly network-based intrusion detection system using a reliable hybrid artificial bee colony and AdaBoost algorithms, Journal of King Saud University-Computer and Information Sciences 31.4 (2019) pp. 541-553

[20] Aljawarneh, Shadi, Monther Aldwairi, and Muneer Bani Yassein., Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model, Journal of Computational Science 25 (2018) pp. 152-160

[21] Nkenyereye, Lewis, Bayu Adhi Tama, and Sunghoon Lim, A Stacking-Based Deep Neural Network Approach for Effective Network Anomaly Detection, Cmc-Computers Materials & Continua 66.2 (2021) pp. 2217-2227

[22] Aggarwal, Preeti, and Sudhir Kumar Sharma, Analysis of KDD dataset attributes-class wise for intrusion detection, Procedia Computer Science 57 (2015) pp. 842-851

[23] Hosseini, Soodeh, and Hossein Seilani, Anomaly process detection using negative selection algorithm and classification techniques, Evolving Systems (2019) pp. 1-10

[24] Mahdavifar, Samaneh, and Ali A. Ghorbani, Application of deep learning to cybersecurity: A survey, Neurocomputing 347 (2019) pp. 149-176

[25] Zhiqiang, Liu, et al., A Three-Layer Architecture for Intelligent Intrusion Detection Using Deep Learning, Proceedings of Fifth International Congress on Information and Communication Technology. Springer, Singapore, 2021, pp. 245-255

[26] Sumaiya Thaseen, I., et al., An integrated intrusion detection system using correlation-based attribute selection and artificial neural network, Transactions on Emerging Telecommunications Technologies 32.2 (2021) e4014

[27] Altunay, H. C., Albayrak, Z., Özalp, A. N., & Çakmak, M. (2021, June) Analysis of anomaly detection approaches performed through deep learning methods in SCADA systems. In 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA) (pp. 1-6) IEEE

[28] Kasongo, Sydney M., and Yanxia Sun, Performance Analysis of Intrusion Detection Systems Using a Feature Selection Method on the UNSW-NB15 Dataset, Journal of Big Data 7.1 (2020) pp. 1-20

[29] Tang, Chaofei, Nurbol Luktarhan, and Yuxin Zhao, SAAE-DNN: Deep Learning Method on Intrusion Detection, Symmetry 12.10 (2020) pp. 1695

[30] Gwon, Hyeokmin, et al., Network Intrusion Detection based on LSTM and Feature Embedding., arXiv preprint arXiv: 1911.11552 (2019)

[31] Ring, Markus, et al., A survey of network-based intrusion detection data sets. Computers & Security 86 (2019) pp. 147-167

[32] Umamaheswari, K., Subbiah Janakiraman, and K. Chandraprabha., Multilevel Hybrid Firefly-Based Bayesian Classifier for Intrusion Detection in Huge Imbalanced Data, Journal of Testing and Evaluation 49.1 (2021)

[33] Milenkoski, Aleksandar, et al., Evaluating computer intrusion detection systems: A survey of common practices, ACM Computing Surveys (CSUR) 48.1 (2015) pp. 1-41

[34] Sabaz, F., & Celik, Y. (2018) Systematic Literature Review on Security Vulnerabilities and Attack Methods in Web Services. International Conference on Advanced Technologies, Computer Engineering and Science (pp. 821-825)

[35] Akter, M., Dip, G. D., Mira, M. S., Hamid, M. A., & Mridha, M. F., Construing attacks of internet of things (IoT) and a prehensile intrusion detection system for anomaly detection using deep learning approach, Springer In International Conference on Innovative Computing and Communications, 2020, pp. 427-438

[36] Ferrag, Mohamed Amine, et al., Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study, Journal of Information Security and Applications 50 (2020) pp. 102419

[37] Buczak, Anna L., and Erhan Guven, A survey of data mining and machine learning methods for cyber security intrusion detection, IEEE Communications surveys & tutorials 18.2 (2015) pp. 1153-1176

[38] Mendez Mena, Diego, and Baijian Yang., Decentralized Actionable Cyber Threat Intelligence for Networks and the Internet of Things, IoT 2.1 (2021) pp. 1-16

[39] Issa, A. S., & Albayrak, Z. CLSTMNet: A Deep Learning Model for Intrusion Detection, Journal of Physics: Conference Series (2021)

[40] Pham, Ngoc Tu, et al., Improving performance of intrusion detection system using ensemble methods and feature selection, Proceedings of the Australasian Computer Science Week Multiconference. 2018, pp. 1-6

[41] Al-Jarrah, O. Y., et al., Machine-learning-based feature selection techniques for large-scale network intrusion detection, 2014 IEEE 34th international conference on distributed computing systems workshops (ICDCSW) IEEE, 2014

[42] Said A. A., Çakmak M. & Albayrak, Z., 6th International Conference on Smart City Applications (SCA2021) (2021) (pp. 1133-1140)

[43] Karimipour, Hadis, and Henry Leung, Relaxation-based anomaly detection in cyber-physical systems using ensemble kalman filter, IET Cyber-Physical Systems: Theory & Applications 5.1 (2020) pp. 49-58

[44] Patil, Tina R., MSSS Performance analysis of naive bayes and J48 classification algorithm for data classification, Journal of Computer Science and Applications 6.2 (2013)

[45] Bhati, Nitesh Singh, and Manju Khari, A Survey on Hybrid Intrusion Detection Techniques, Research in Intelligent and Computing in Engineering. Springer, Singapore, 2021, pp. 815-825

[46] Azzaoui, Hanane, et al., Developing new deep-learning model to enhance network intrusion classification, Evolving Systems (2021) pp. 1-9

Creep and Reliability Prediction of a Fan-Out WLP Influenced by the Visco-Plastic Properties of the Solder

Ramiro Sebastian Vargas Cruz[1] and Viktor Gonda[2]

[1]Óbuda University, Doctoral School on Materials Sciences and Technologies,
Bécsi út 96/b, 1034 Budapest, Hungary
ramiro.vargas@stud.uni-obuda.hu

[2]Óbuda University, Donát Bánki Faculty Mechanical and of Safety Engineering,
Népszínház u. 8, 1081 Budapest, Hungary
gonda.viktor@bgk.uni-obuda.hu

Abstract: Solder joint reliability is critical in the design of advanced microelectronic packaging. Predictions of reliability by thermo-mechanical simulations can accelerate the evaluations of advanced packaging and the introduction of novel solder materials. In this work, a finite element model of a thermally loaded Fan-Out Wafer Level Package (FO-WLP) was built and analyzed focusing on the creep behavior of the solder balls and the consequent effect on the reliability of the package. The lead-free soldering materials in the analyses were either of a widely used SAC305, or novel doped SAC solders as SAC-R, SAC-Q and InnoLot. Visco-plastic (Anand creep) properties for the solders were defined as study parameters, where 6 variations were used for the described SAC305, and further 3 sets for the doped SAC solders, respectively. Identifying a stress concentration at the sharp bond pad edges by modeling ideal geometries, a refined geometry was introduced and evaluated. Simulations for a 3-cycle thermal load were conducted, and results were collected and analyzed for Creep Strain and Strain Energies in critical positions in the solder, and reliability prediction was performed based on Morrow's model. Results show the benefit of the refined compositions of Doped SAC solders on the mechanical behavior and improved reliability.

Keywords: Reliability; Creep behavior; lead-free Solder

1 Introduction

The international pursuit of an environment free of hazardous substances and compliance with international regulations has undoubtedly affected how technology developed during the last years. According to the application, the allowed percentage of lead, cadmium, and other composites in electronic devices has been limited [1]. One of the most prevalent materials in microelectronic interconnects is the solder alloy, which mechanically and electronically connects the components to the substrate [2]. Therefore, lead-free tin (Sn) based soldering materials have taken the market based on their similar performance to the eutectic tin-lead (SnPb) solder [3].

Simultaneously, the relentless technological development has looked for more compact electronic devices. Integrated Circuits (IC) are part of several applications, like the automotive industry, where the components undergo harsh environments and high temperatures [4]. Therefore, it is imperative to assure a reasonable lifetime of every element of the electronic board. This work aims to analyze the structural behaviour of an advanced microelectronic package by finite element simulations. A thermally loaded Fan-Out Wafer Level Package (FO-WLP) was modelled and analyzed focusing on the creep behavior of the solder balls and the consequent effect on the reliability of the package. The lead-free soldering materials in the analyses were either of a widely used SAC305, or novel doped SAC solders as SAC-R, SAC-Q and InnoLot. Visco-plastic (Anand creep) properties for the solders were defined as study parameters, where 6 variations were used for the described SAC305, and further 3 sets for the doped SAC solders, respectively. Simulations for a 3-cycle thermal load were conducted, and results were collected and analyzed for Creep Strain and Strain Energies in critical positions in the solder, and reliability prediction was performed. The outline of the paper is as follows: the Section 2 presents a concise introduction of electronic packaging. Subsequently, mathematical models that describe creep behavior and reliability for solders are presented. The section details the modelling approach, followed by a thorough analysis in the results and discussion sections.

2 Fan-Out Wafer Level Packaging (FO-WLP)

The electronic industry has been considered a crucial sector for technology development since the first working transistor was announced by AT&T (American Telephone and Telegraph) at Bell Laboratories. Hence, silicon-based transistors later developed by Texas Instruments resulted in a breakthrough with more reliable switching states [5]. Since 1970, Integrated Circuits (IC) have been part of the advancement in technologies and the computing industry. The chronological evolution of electronic packaging is detailed in Figure 1. The need for more efficient ICs constantly motivates the industry to increase the number of input/output (I/O) pins, while miniaturizing ICs.

During the last decade, Wafer Level Packaging (WLP) has been utilized to rapidly increase the number of I/O while reducing ICs size. Lau et al. [7] presented a concise review of the materials and trends regarding Fan-In and Fan-Out – WLP. Some of the advantages of WLP are a substrate-less package, lower thermal resistance, and higher performance due to shorter interconnects [8].

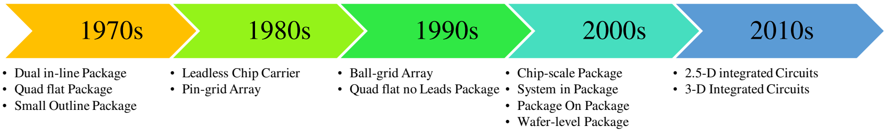


Figure 1
Evolution of semiconductor packaging [6]

Two basic Fan-Out WLP (FO-WLP) structures can be identified: Mold first and Redistribution Layer (RDL) first, based on the process flow (see Figure 2).

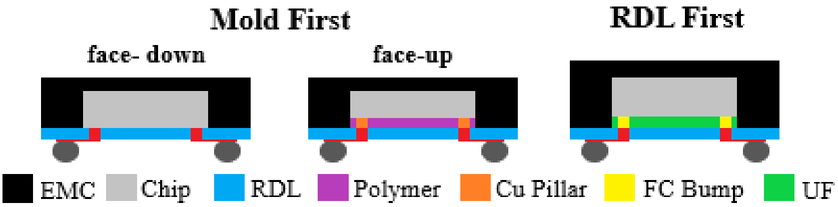


Figure 2
FO WLP Structures [9]
EMC: Epoxy Molding Compound, Cu: copper, FC Bump: Flip Chip Bump, UF: Underfill, and RDL: Redistribution Layer

Mold first – FO-WLP is already in mass production. Its advantage is related to the heterogeneous components needed in panels assemblies for energy harvest [9]. In addition, recent studies demonstrate that RLD first – FO-WLP could potentially replace some of the 3D ICs [10]. However, RDL first faces challenge to enhance wafer warpage and RDL scaling [11].

3 Solder Joint Modeling and Reliability

Reliability engineering has become a core part of all branches of industry. It consists of three main concepts (design for reliability DFR, reliability test & data analysis, and failure analysis) to effectively optimize the sources to predict failure [12]. The initial stage of reliability engineering (DFR) comprises the simulation procedure; once the optimal materials and measurements are manufactured in samples, reliability tests are conducted. The resulting data is then used for lifetime prediction. Finally, the failure analysis aims to understand the root cause based on faulty samples from the second stage. Recently, some standard tests for solder joint reliability have been identified. These tests usually imply endurance to changes in temperature, specific mechanical loads, and drop tests. Since creep behavior involves temperature, stress, and time, it has been implemented in several Finite Element Method (FEM) software packages [13]. The most common creep and reliability models are summarized in this section.

3.1 Anand Creep Model for Soldering Materials

Creep is time-dependent plastic deformation significant above half the absolute melting point [14]. Creep comprises three stages: primary, secondary, and tertiary. Along the primary stage (strain–hardening), the strain rate decreases as the hardening increases. Then, strain rate becomes constant; this is the steady-state creep. Finally, the strain rate rises exponentially in the tertiary stage, leading to a material fracture [15].

Mathematically, steady state creep behavior can be described by Norton Power Law [16], Garofalo-Arrhenius constitutive model [17], or Anand's constitutive model [18]. Norton's description is a basic approach that is refined in several subsequent work. The Garofalo-Arrhenius model has better accuracy [19] and extensively implemented in finite element software [20].

Anand [18] proposed two evolution equations from the flow equation (1) based on earlier theories suggested by Lee and Zaverl [21]. The primary purpose of this research was a better analysis of the deformation of metal alloys at elevated temperatures above $0.5T_m$ (absolute melting temperature). After a concice mathematical analysis on viscoplastic deformation at elevated temperatures, equations (2), (3), and (4) were finally proposed [22].

Flow equation:

$$\dot{\varepsilon}_p = A \cdot \exp\left(-\frac{Q}{RT}\right) \cdot \left(\sinh\left(\xi \frac{\sigma}{s}\right)\right)^{1/m} \tag{1}$$

where A is the pre-exponential factor, R is the universal gas constant, ξ is the multiplier of stress, s is the deformation resistance, and m is the strain rate sensitivity of stress.

Evolution equations:

$$B = 1 - \frac{s}{s^*} \tag{2}$$

$$\dot{s} = \{h_0 \cdot |B|^a \cdot \mathrm{sgn}(B)\}\dot{\varepsilon}_p \dot{\mathrm{s}} = \{\mathrm{h_0}|\mathrm{B}|^{\mathrm{a}}\mathrm{sgn(B)}\}\dot{\varepsilon_{\mathrm{p}}} \tag{3}$$

$$s^* = \hat{s} \cdot \left(\frac{\dot{\varepsilon}_p}{A} \exp\left(\frac{Q}{RT}\right)\right)^n \tag{4}$$

where, s^* is the saturation resistance, h_0 is the hardening constant, a is the strain rate sensitivity of hardening, $\hat{s}$ is the deformation resistance saturation coefficient, and n is the strain rate sensitivity of saturation.

Anand model involves nine material parameters that can be determined following Brown's procedure [23]. Brown suggests at least two sets of three strain rate jump tests performed at different temperatures.

3.2 Failure Prediction Models

In ball grid array type packages, the Distance to Neutral Point (DNP) concept is widely employed to select critical solder joints [24]. The effectiveness of the DNP approach has been intensely discussed. Lau [25] stresses that DNP lifetime approximation presents some limitations. For instance, it was proven to be valid for packaging without underfill; however, it has been incorrectly applied on packaging with underfill [26]. The lifetime of the critical solder joint can be estimated by fatigue modells, such as the Coffin-Manson model based on strains, or the energy based Morrow's model among others.

3.2.1 Coffin-Manson Model

The Coffin-Manson model is strain-based analysis for low-cycle fatigue [27], where the equation is given as:

$$N_f \left(\Delta\varepsilon_p\right)^n = C \tag{5}$$

where, N_f is the predicted number of cycles to failure, n is the empirical constant, $\Delta\varepsilon_p$ is the inelastic strain range, and C is the proportionality factor/fatigue ductility coefficient.

According to Norris et al. [28], constant n was observed to be 2 for most metals. This mathematical approach has been widely studied. Zubelewicz et al. [29] stressed that the Coffin-Manson law does not fit experimental data well for high strain rates.

3.2.2 Morrow's Model

Morrow [30] proposed an energy-based equation to predict the number of cycles to failure:

$$N_f{}^{n'} W_p = A \tag{6}$$

where, n' is the fatigue exponent, A is the material ductility coefficient, and W_p is the inelastic strain energy density.

4 Simulation Procedure

This section thoroughly details the parameters for the electronic packaging simulations. First, the model's geometry, boundary conditions, and finally, the material properties are described. The simulations were carried out in Marc Mentat finite element software.

4.1 Geometry

For the finite element modeling, a 167GJJ Package from Texas Instruments [31] was selected. This FO-WLP package consists of a fine pitch Ball Grid Array (BGA). The main dimensions are shown in Figure 3.

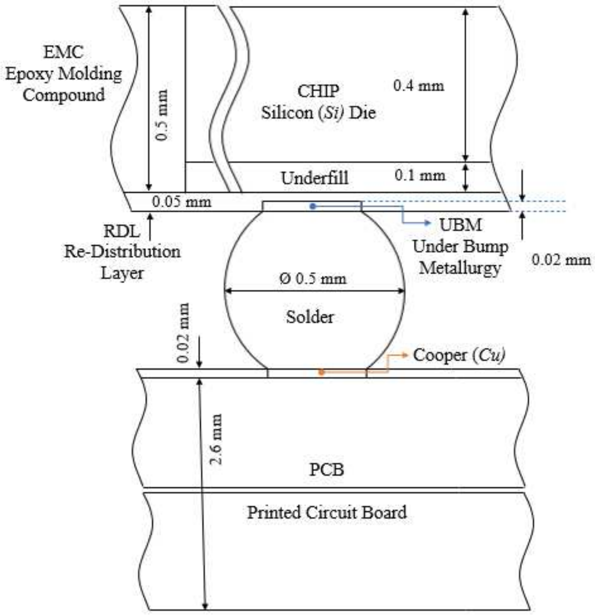


Figure 3
Main dimensions of the FO-WLP package section

The bottom copper layer was initially modeled with sharp pad edges considering an ideal geometry in previous works [32], [33]. However, during the processing, this edge may become blunt in real geometries. Therefore, a second variant is modeled using a fillet on the bottom copper layer. For simplicity, the model with the sharp-edged pad will be referred to as “squared” while the model with blunt-edged pads will be referred to as “rounded” (see Figure 4). Additionally, due to the complexity of the mesh, a mask layer was included in the squared model, whereas the same mask layer was excluded in the rounded model.

Two dimensional modelling was created to optimize the computing resources. Therefore, only half of the cross-section of the package was modeled. The complexity of the mesh utilized is not visible in Figure 4, the modeled layout is shown with the incorporated materials. The squared model was created using ruled mesh, while the rounded model was created using mesh seeds due to the fillet on the copper pad; the complete mesh contains over 200 000 elements and 115 000 nodes.

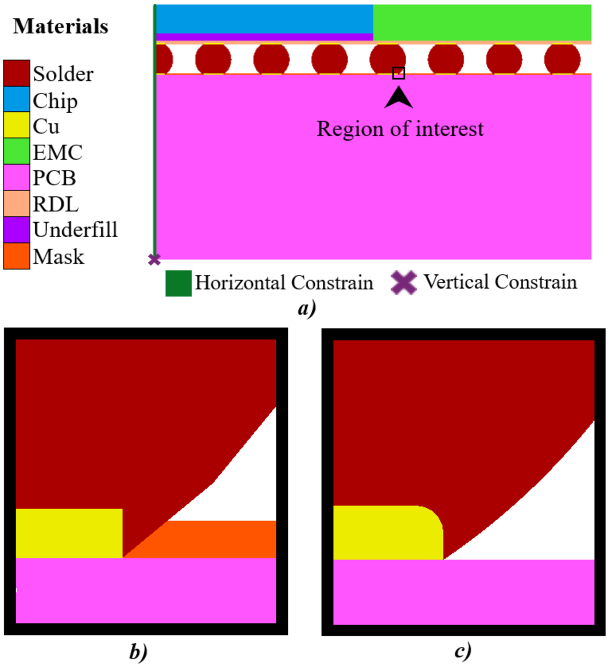


Figure 4

The modelled section: *a)* Materials description and boundary conditions, *b)* Augmented view of the squared Cu pad profile, and *c)* Augmented view of the rounded Cu pad profile

4.2 Initial and Boundary Conditions

A temperature of -40°C was set for all nodes for initial thermal conditions. For mechanical boundary conditions, mirror symmetry was considered by constraining the horizontal displacement on the model's left side. Constraints to vertical displacement were applied to the bottom left node for a unique solution (Figure 4). No extra mechanical loads were applied since the coefficient of thermal expansion mismatch will produce mechanical stress in the solder.

A cyclic thermal loading was applied to the modeled section. Previous studies on failure prediction have shown that a stable plastic work density is reached within three thermal cycles [34]. An average of strain and strain energy density can be computed, taking values at the second and third thermal cycles. Hence, a three-hour thermal load was applied to all the nodes in the model. The temperature ranged from -40°C to 125°C, as shown in Figure 5. This temperature range is in accordance with the Joint Electron Device Engineering Council (JEDEC) standards [35], and the maximum temperature surpasses $0.5T_m$. The repetitive cycle starts at -40°C, followed by a ramp-up that reaches 125°C within 15 minutes (heating stage).

Then, the temperature remains constant for the next 15 minutes, followed by a ramp-down that reaches -40°C in 15 minutes (cooling stage). Finally, the temperature is held for 15 minutes before the start of the next cycle.

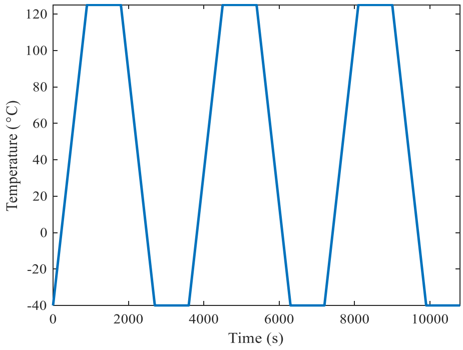


Figure 5

Thermal Load

In the finite element solver, a multi-criteria time stepping was selected for stable solutions with refined target step size, assuring the sufficiently large number of steps per cycle to maintain the accuracy of the calculations. Due to the fine mesh utilized for accurate results and small timesteps, the simulation process took about 8 hours for the squared profile and 6 hours for the rounded profile.

4.3 Material Properties

The material properties as Young's Modulus (E), Poisson's Ratio (υ), and Coefficient of Thermal Expansion (CTE) are summarized in Table 1 for the materials excluding the solder. These properties were considered time and temperature-independent, except for the Glass transition temperature (Tg) for certain materials. Under Bump Metallization (UBM) typically consist of Copper (Cu) coated with a noble metal to avoid corrosion, *e.g.,* Gold (Au) [36]. For simplification, the pad was modeled entirely of copper.

Table 1

Material Properties [37]

Material	***E* (GPa)**		***υ***	**CTE (ppm)**		**Tg* (°C)**
	T≤Tg	T>Tg		T≤Tg	T>Tg	
Chip	313		0.30	2.8		
Cu	117		0.34	17		
EMC	18.5	1.2	0.30	9	18	163
PCB	25		0.11	15		
RDL	0.92	0.1	0.30	80	227	205

Underfill	3.8	0.125	0.30	44	119	141
Mask	2.4	0.23	0.30	60	161	100

*Glass Transition Temperature

Table 2

Chemical composition (wt%) of the selected SAC solders

Solder	**Sn**	**Ag**	**Cu**	**Bi**	**Ni**	**Sb**
SAC-R [38]	96.62	0.00	0.92	2.46	0.00	0.00
SAC-Q [38]	92.77	3.41	0.52	3.30	0.00	0.00
InnoLot [38]	90.95	3.80	0.70	3.00	0.15	1.40
SAC305 [39]	95 – 96	3.8 – 4.2	0.3 – 0.7	0.00	0.00	0.00

Table 3

Anand Properties for the solder materials

	S_0	A	ξ	m	h_0	s	n	a	Q
	MPa	s^{-1}	-	-	MPa	MPa	-	-	$J\cdot mol^{-1}$
SAC305 from different experimental works									
Alam [41]	6.5	3700	4	0.47	70000	7.72	0.0315	1.9	95616.75
Basit [42]	21	3501	4	0.25	$180\cdot10^3$	30.2	0.01	1.78	77491.14
Herk. [43]	1.066	$1.43\cdot10^8$	1.472	0.14	5023.9	20.29	0.032	1.12	88581.38
Janz [44]	45.9	$5.87\cdot10^6$	2	0.09	9350	58.3	0.015	1.5	62026.17
Lall [45]	32.39	1100	6	0.39	174130	67.7	0.0008	1.75	33258
Mysore [46]	2.15	17.994	0.35	0.15	1525.98	2.53	0.028	1.69	82895.56
Doped SAC Solders									
SACQ [47]	0.405	$2.45\cdot10^8$	0.068	0.36	3521.56	0.638	0.0056	1.243	112313.8
SACR [38]	34.72	1000	6	0.15	145640	71.71	0.001	1.55	92290.52
InnoLot [38]	32.42	25000	7	0.35	88875	56.76	0.0097	1.45	89313.9

Four different lead-free Sn-Ag-Cu (Tin-Silver-Coper) type soldering materials were considered in this work; SAC305 and three doped SAC solders (SACX) (see Table 2). Temperature-dependent *E* and CTE were set for SAC305 (see Figure 6).

For creep modeling, several sets of Anand Parameters are available for SAC305 in the literature (Table 3), and Anand parameters for the SACX solders are summarized as well.

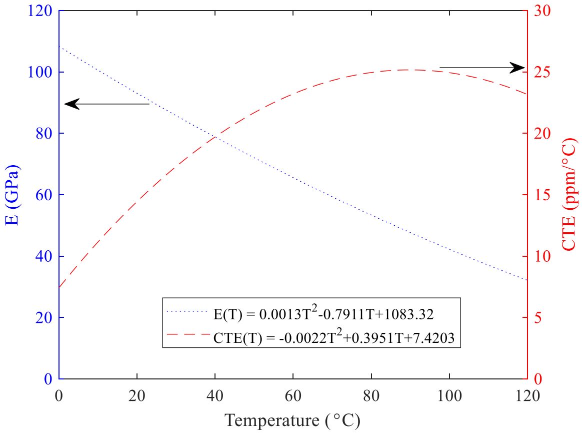


Figure 6
Temperature-dependent elastic modulus and CTE for SAC305 [40]

4.4 Overview of the Studied Cases

Alltogether, 11 cases are defined and summarized in Table 4. The main parameters were the material properties for the solder, and the solder pad geometry. The initial step was to compare the different sets of Anand parameters for SAC305 (Case 01 - 06). Next, a comparison between SAC305 and Doped SAC solders was carried out (Case 07 - 09). Once the results were analyzed, selected material properties were simulated with the round pad profile (Case 10 and 11).

Table 4
Summary of the main studied parameters

Case	Material	Edge Shape		Case	Material	Edge Shape	
		Sq.	Ro.			Sq.	Ro.
SAC305 Authors Comparison				Doped SAC Solders			
01	Alam	×		**07**	SACQ	×	
02	Basit	×		**08**	SACR	×	
03	Herkommer	×		**09**	InnoLot	×	
04	Janz	×		Edge shape comparison purposes			
05	Lall	×		**10**	Basit		×
06	Mysore	×		**11**	SACQ		×

Sq. = Squared, Ro = Rounded

5 Results and Discussion

The contour band graph of the Total Equivalent of Creep Strain (TECS) at the end of the simulation is shown in Figure 7. The node with the highest TECS is located on the outer bottom side of the solder ball. This result agrees with experimental data presented in [13], where in packaging including UBM cracking initiates in the bottom corners, while in packages without UBM cracking initiates in the upper corners. The magnified image in the bottom part of Figure 7 shows a significant change of the critical node location and the maximum value reached (0.22 for the squared profile and 0.09 for the rounded profile). The second most critical point occurred in the middle solder ball (4th from right to left on Figure 7). This result can be explained by the high number of materials in the vicinity of the solder ball. In addition, a mismatch of CTE increases the stress generated in the solder ball.

The levels of TECS in the upper side of the solder balls display a high concentration of strain under the EMC layer (see Figure 4 and Figure 7). Compared to experimental results, simulations on Wafer Level Chip Scale Package (WLCSP) support this phenomenon since solder crack initiated in the outer upper part of the solder ball [48].

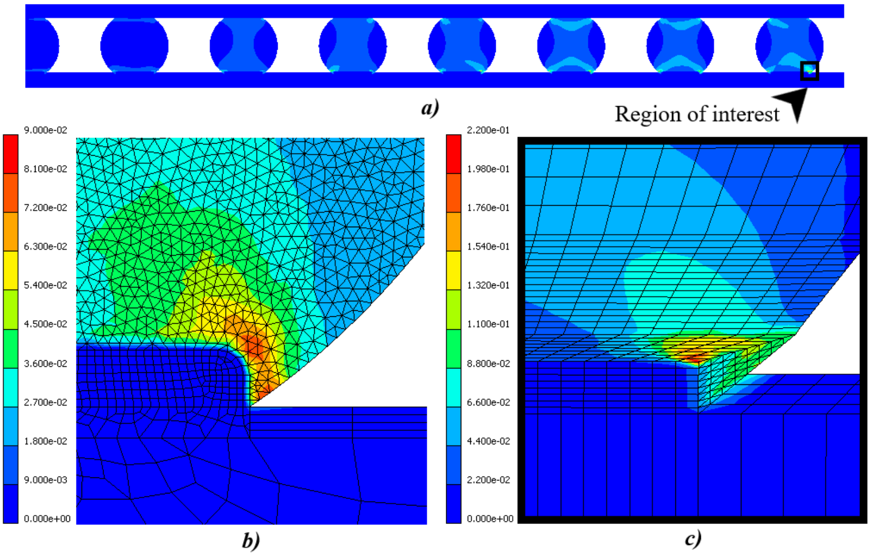


Figure 7

Total Equivalent of Creep Strain at the end of the three cycles: *a)* TECS distribution along the solder balls (cropped image), Augmented contour band graph of TECS distribution at the vicinity of the critical node *b)* rounded Cu pad profile and, *c)* squared Cu pad profile

5.1 Comparison of SAC305 Creep Parameters

TECS vs. time curves for the six sets of SAC305 Anand Parameters are plotted in Figure 8. The six replicates (study case 01 to 06) agreed on the most critical node location (outer bottom side of the solder ball as shown in Figure 7). There is a significant difference between the models of Mysore and Janz as compared to the rest of the Authors. Parameters of Alam, Basit, and Lall results in strains that are considerably similar. In terms of simulation time, the Basit parameters took significantly less time than Alam's and Lall's.

Although the six authors utilized the same composition, different experiments were followed to obtain Anand parameters. Shear stress and normal stress were measured and tabulated with shear strain and normal strain, respectively. The results displayed in Figure 8 cannot by established if the parameters are correct, but it can be a valuable tool for simulations where the time needs to be minimized. In such a case, Basit values approximate similar results to the remaining sets of Anand parameters while taking less time.

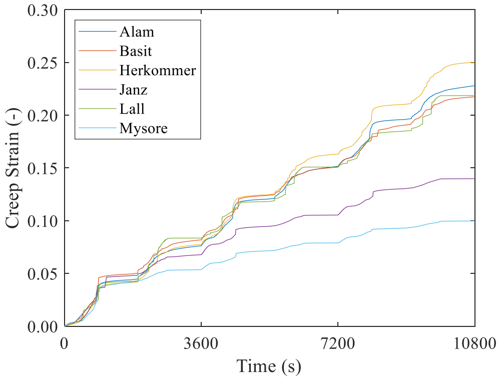


Figure 8
Total Equivalent of Creep Strain at the critical node – SAC305. Squared Profile

5.2 Comparison of SAC305 to Doped SAC Solders

Since the simulation using Anand parameters from Basit took less time to run, a second comparison was carried out between SAC305 (Basit) and Doped SAC solders. Creep Strain Energy Density (CSED) data was collected from the four replicates (case 02, 07 - 09). CSED vs. time curves are presented in Figure 9. According to Che and Pang [34]**,** an approximation of the average of CSED can be computed taking values from the second and third thermal cycles. The obtained values were as follow: SAC305 = 0.2651 MPa; SACQ = 0.1627 MPa; InnoLot = 0.2550 MPa and SACR = 0.2019 MPa.

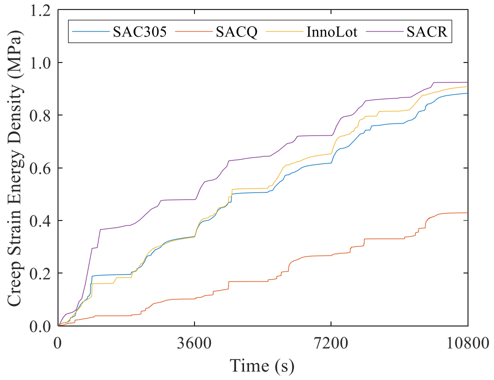


Figure 9

Creep Strain Energy Density SAC305 vs. Doped SAC Solders. Squared Profile

Based on Morrow's Model, presented in Section 3.2.2, the number of cycles to failure is inversely proportional to the average inelastic strain energy density. Therefore, in terms of CSED, SACQ would ideally perform longer than other Doped SAC solders and SAC305.

5.3 Effect of the Bond Pad Geometries

A comparison regarding the copper pad profile was carried out. Since the geometry of the modified model become too complex, the mask layer was neglected following similar study cases [49]. The fillet radius was assumed to be 20% of the copper pad width based on X-ray inspections taken from Lau [13].

Creeps strain curves (Equivalent of Creep Strain – ECS and TECS) from the critical node are presented in Figure 10. Like the squared profile, the location of the critical node remained the same.

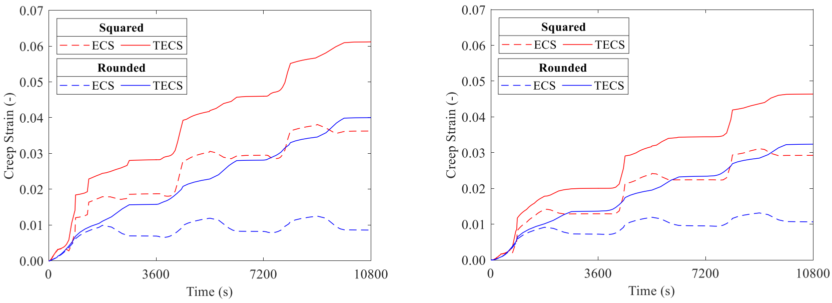
The variation of change in profile within the same composition and between compositions are presented in Table 5. The column variations (Table 5) represent the difference of TECS between materials, while the row variations represent the variation due to the change in profile shape. It is clear to notice that in both compositions, the decrement due to shape is nearly 30%. On the other hand, a rounded profile accentuates the TECS difference between materials by 4.97%.

Regarding to Creep Strain Energy Density (CSED) and Total Strain Energy Density (TSED), curves are presented in Figure 11. In the case of SACQ, TSED for the squared profile presents several unstable peaks, and the general TSED values are not greater than those of CSED. On the other hand, TSED values are constantly greater than the CSED values for the rounded profile, and the curves follow a more stable pattern. In both cases, SAC305 and SACQ squared profile; the peaks show a time dependency and relaxation.

Table 5

TECS final value variation

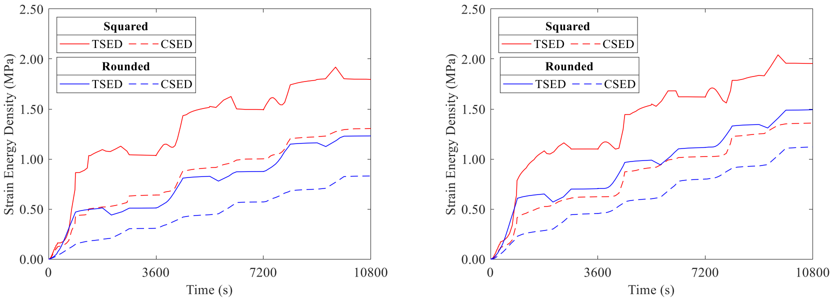
	SAC305	**SACQ**	**variation**
Squared	0.0611	0.0463	**24.22%**
Rounded	0.0400	0.0323	**19.25%**
variation	**34.53%**	**30.23%**	



a) SAC305 b) SACQ

Figure 10

Creep Strain Curves



a) SAC305 b) SACQ

Figure 11

Strain Energy Density curves

CSED final values are summarized in Table 6. SAC305 reduced 31.45% due to a change in shape, while SACQ reduced 23.61%. It is interesting to see that the initial difference between materials was 8.11% using a squared profile. A change in profile increased the difference between materials to 17.54%. From Table 5 and Table 6, a change in the copper pad profile accentuates the variation of creep performance between materials.

Table 6

TSED final value variation

	SAC305	**SACQ**	**variation**
Squared	1.7960	1.9546	**8.11%**
Rounded	1.2311	1.4930	**17.54%**
variation	**31.45%**	**23.61%**	

5.4 Reliability Prediction Analysis

Following Che and Pang's procedure [34] (subtraction between final values of the third and second cycle), values of the approximation of average strain and energy variation were computed (see Table 7). It should be noted that all values decreased due to a change in copper pad geometry. High values can be observed for the squared profile with a significant reduction for the rounded profile. However, a more stable curve regarding TSED in SACQ (Case 11) increased inelastic strain energy density.

Table 7

Approximation of average of creep strain and energy density

	SAC305		**SACQ**	
	W_p	$\Delta\varepsilon_p$	W_p	$\Delta\varepsilon_p$
Squared	0.3028	0.0152	0.3347	0.0120
Rounded	0.3558	0.0119	0.3770	0.0090

Recent studies have reported Morrow and Coffin-Manson constants for estimation of the number of cycles to failure. It has been shown that the temperature, frequency, and aging may potentially affect the constants [50]. Morrow and Coffin-Manson constants for SAC305 and SACQ aging dependent have been reported in the last two years [51]. Nevertheless, temperature, and frequency dependency is needed for an accurate estimation using Table 7.

Conclusions

The reliability of the solder ball can compromise the functionality of the entire Integrated Circuit. Repetitive thermal loading of a complex FO-WLP package was modeled, varying the copper profile and solder ball material viscoplastic properties. Based on the results, the following conclusions can be drawn:

1) From the SAC305 comparison, six different sets of Anand Parameters were tested using simulation replicates. Not only the creep behavior was different, but also the simulation time. Basit parameters are suitable for simulation since it takes the minimum time, and the results do not significantly differ from the rest.

2) Since reliability models are inversely related to inelastic strain energy, SACQ presented the most promising working time, followed by SACR, Innolot, and SAC305 in that order.

3) A change in the copper pad profile shape (squared to rounded) shows a stress reduction and, therefore, more stable creep curves. Additionally, it accentuates the difference of creep values between materials by nearly 16% regarding creep strain values.

References

[1] European Parliament; The Council of The European Union, "Directive 2011/65/Eu of The European Parliament and of The CouncIL of 8 June 2011 on the restriction of the use of certain hazardous substances in electrical and electronic equipment (recast)," *Official Journal of the European Union*, 2011

[2] P. D. Sonawwanay and V. K. Bupesh Raja, "Eco-friendly Soldering Technique," in *Techno-Societal 2020*, Cham: Springer International Publishing, 2021, pp. 761-766

[3] J. Shepherd, "Lead Restrictions and Other Regulatory Influences on the Electronics Industry," in *Lead-Free Soldering*, J. Bath, Ed. Boston, MA: Springer US, 2007, pp. 5-19

[4] M. Traub, A. Maier, and K. L. Barbehon, "Future Automotive Architecture and the Impact of IT Trends," *IEEE Softw.*, Vol. 34, No. 3, pp. 27-32, May 2017, doi: 10.1109/MS.2017.69

[5] M. Riordan, "The lost history of the transistor," *IEEE Spectr.*, Vol. 41, No. 5, pp. 44-49, May 2004, doi: 10.1109/MSPEC.2004.1296014

[6] AnySilicon, "Semiconductor Packaging History and Trends," 2016. https://anysilicon.com/semiconductor-packaging-history-trends/

[7] J. H. Lau *et al.*, "Design, Materials, Process, Fabrication, and Reliability of Fan-Out Wafer-Level Packaging," *IEEE Trans. Components, Packag. Manuf. Technol.*, Vol. 8, No. 6, pp. 991-1002, Jun. 2018, doi: 10.1109/TCPMT.2018.2814595

[8] T. Braun *et al.*, "Opportunities of Fan-out Wafer Level Packaging (FOWLP) for RF applications," in *2016 IEEE 16th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)*, Jan. 2016, pp. 35-37, doi: 10.1109/SIRF.2016.7445461

[9] T. Braun *et al.*, "Fan-Out Wafer and Panel Level Packaging as Packaging Platform for Heterogeneous Integration," *Micromachines*, Vol. 10, No. 5, p. 342, May 2019, doi: 10.3390/mi10050342

[10] L. T. Guan, C. K. Fai, and D. H. S. Wee, "FOWLP electrical performances," in *2016 IEEE 18th Electronics Packaging Technology Conference (EPTC)*, Nov. 2016, Vol. 3, pp. 79-84, doi: 10.1109/EPTC.2016.7861447

[11] V. S. Rao *et al.*, "Development of High Density Fan Out Wafer Level Package (HD FOWLP) with Multi-layer Fine Pitch RDL for Mobile Applications," in *2016 IEEE 66th Electronic Components and Technology Conference (ECTC)*, May 2016, pp. 1522-1529, doi: 10.1109/ECTC.2016.203

[12] A. Birolini, *Reliability engineering: theory and practice*. Springer Science & Business Media, 2013

[13] J. H. Lau, "State of the Art of Lead-Free Solder Joint Reliability," *J. Electron. Packag.*, Vol. 143, No. 2, Jun. 2021, doi: 10.1115/1.4048037

[14] M. E. Kassner, *Fundamentals of Creep in Metals and Alloys*. Elsevier, 2015

[15] G. Dieter, "Mechanical Metallurgy." p. 615, 1961

[16] F. H. Norton, *The creep of steel at high temperatures*, No. 35, McGraw-Hill Book Company, Incorporated, 1929

[17] F. Garofalo and D. B. Butrymowicz, "Fundamentals of Creep and Creep-Rupture in Metals," *Phys. Today*, Vol. 19, No. 5, pp. 100-102, May 1966, doi: 10.1063/1.3048224.

[18] L. Anand, "Constitutive Equations for the Rate-Dependent Deformation of Metals at Elevated Temperatures," *J. Eng. Mater. Technol.*, Vol. 104, No. 1, p. 12, 1982, doi: 10.1115/1.3225028

[19] J. H. Lau, "Creep of 96.5Sn3.5Ag Solder Interconnects," *Solder. Surf. Mt. Technol.*, Vol. 5, No. 3, pp. 45-52, Mar. 1993, doi: 10.1108/eb037839

[20] J. A. Depiver, S. Mallik, and E. H. Amalu, "Effective Solder for Improved Thermo-Mechanical Reliability of Solder Joints in a Ball Grid Array (BGA) Soldered on Printed Circuit Board (PCB)," *J. Electron. Mater.*, Vol. 50, No. 1, pp. 263-282, Jan. 2021, doi: 10.1007/s11664-020-08525-9

[21] D. Lee and F. Zaverl Jr, "A generalized strain rate dependent constitutive equation for anisotropic metals," *Acta Metall.*, Vol. 26, No. 11, pp. 1771-1780, 1978

[22] L. Anand, "Constitutive equations for hot-working of metals," *Int. J. Plast.*, Vol. 1, No. 3, pp. 213-231, 1985, doi: 10.1016/0749-6419(85)90004-X

[23] S. B. Brown, "An Internal Variable Constitutive Model for the Thixotropic Behavior of Metal Semi-Solid Slurries," *Materials Science Seminar on Intelligent Processing of Materials*, Vol. 5. pp. 95-130, 1989

[24] T. C. Lui and B. N. Muthuraman, "Relibability assessment of wafer level chip scale package (WLCSP) based on distance-to-neutral point (DNP)," in *2016 22nd International Workshop on Thermal Investigations of ICs and Systems (THERMINIC)*, Sep. 2016, pp. 268-271, doi: 10.1109/THERMINIC.2016.7749063

[25] J. H. Lau, "The Roles of DNP (Distance to Neutral point) on Solder Joint

Reliability of Area Array Assemblies," *Solder. Surf. Mt. Technol.*, Vol. 9, No. 2, pp. 58-60, Dec. 1997, doi: 10.1108/09540919710800674

[26] W. Dauksher and W. S. Burton, "An examination of the applicability of the DNP metric on first level reliability assessments in underfilled electronic packages," *Microelectron. Reliab.*, Vol. 43, No. 12, pp. 2011-2020, Dec. 2003, doi: 10.1016/S0026-2714(03)00219-1

[27] S. S. Manson, "Thermal stress and low-cycle fatigue" 1966

[28] K. C. Norris and A. H. Landzberg, "Reliability of Controlled Collapse Interconnections," *IBM J. Res. Dev.*, Vol. 13, No. 3, pp. 266-271, May 1969, doi: 10.1147/rd.133.0266

[29] A. Zubelewicz, R. Berriche, L. M. Keer, and M. E. Fine, "Life-time prediction of solder materials," *Am. Soc. Mech. Eng.*, Vol. 111, No. September 1989, pp. 179-182, 1988

[30] J. Morrow, "Cyclic Plastic Strain Energy and Fatigue of Metals," in *Internal Friction, Damping, and Cyclic Plasticity*, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959: ASTM International, 1965, pp. 45-87

[31] Texas Instruments Incorporated, "MicroStar BGA, Packaging Reference Guide," no. September. 2000

[32] R. S. Vargas C and V. Gonda, "Sensitivity of the structural behavior of SAC305 interconnects on the variations of creep parameters," May 2021, doi: 10.1109/SACI51354.2021.9465551

[33] R. S. Vargas Cruz and V. Gonda, "Solder joint reliability based on creep strain energy density for SAC305 and doped SAC solders," *MATEC Web Conf.*, Vol. 343, p. 02005, Aug. 2021, doi: 10.1051/matecconf/202134302005

[34] F. X. Che and J. H. L. Pang, "Thermal fatigue reliability analysis for PBGA with Sn-3.8Ag-0.7Cu solder joints," in *Proceedings of 6th Electronics Packaging Technology Conference (EPTC 2004) (IEEE Cat. No.04EX971)*, 2004, pp. 787-792, doi: 10.1109/EPTC.2004.1396715

[35] JEDEC Solid State Technology Association 2000, "JESD22-A104-B," *JEDEC STANDARD*, vol. Temperatur. Arlington, 2009 [Online] Available: http://web.cecs.pdx.edu/~cgshirl/Documents/22a104b Temperature Cycling.pdf

[36] Y. Degani, T. D. Dudderar, and K. L. Tai, "Flip chip packaging of memory chips." Google Patents, Nov. 23, 1999

[37] Z. Chen *et al.*, "Solder Joint Reliability Simulation of Fan-out Wafer Level Package (FOWLP) Considering Viscoelastic Material Properties," *2018 IEEE 20th Electron. Packag. Technol. Conf.*, pp. 573-579, 2019, doi: 10.1109/eptc.2018.8654355

[38] S. Ahmed, M. Basit, J. C. Suhling, and P. Lall, “Characterization of Doped SAC Solder Materials and Determination of Anand Parameters,” in *Vol. 2: Advanced Electronics and Photonics, Packaging Materials and Processing; Advanced Electronics and Photonics: Packaging, Interconnect and Reliability; Fundamentals of Thermal and Fluid Transport in Nano, Micro, and Mini Scales*, Jul. 2015, pp. 1-14, doi: 10.1115/IPACK2015-48624

[39] K. Seelig and D. Suraski, “A COMPARISON OF TIN-SILVER-COPPER LEAD-FREE SOLDER ALLOYS,” *Lead-free Solder Alloy.*, pp. 1-11, 2014

[40] T. T. Nguyen, D. Yu, and S. B. Park, “Characterizing the mechanical properties of actual SAC105, SAC305, and SAC405 solder joints by digital image correlation,” *J. Electron. Mater.*, Vol. 40, No. 6, pp. 1409-1415, 2011, doi: 10.1007/s11664-011-1534-z

[41] M. S. Alam, K. M. R. Hassan, J. C. Suhling, and P. Lall, “High temperature mechanical behavior of SAC and SAC+X lead free solders,” *Proc. - Electron. Components Technol. Conf.*, Vol. 2018-May, pp. 1781-1789, 2018, doi: 10.1109/ECTC.2018.00268

[42] M. Basit, M. Motalab, J. C. Suhling, and P. Lall, “Viscoplastic Constitutive Model for Lead-Free Solder Including Effects of Silver Content, Solidification Profile, and Severe Aging,” in *Volume 2: Advanced Electronics and Photonics, Packaging Materials and Processing; Advanced Electronics and Photonics: Packaging, Interconnect and Reliability; Fundamentals of Thermal and Fluid Transport in Nano, Micro, and Mini Scales*, Jul. 2015, p. V002T01A002, doi: 10.1115/IPACK2015-48619

[43] D. Herkommer, J. Punch, and M. Reid, “Constitutive Modeling of Joint-Scale SAC305 Solder Shear Samples,” *IEEE Trans. Components, Packag. Manuf. Technol.*, Vol. 3, No. 2, pp. 275-281, Feb. 2013, doi: 10.1109/TCPMT.2012.2227481

[44] D. T. Janz, “Reliability of discrete power devices with lead free solder joints.” MS dissertation. Institute for Microsystem Technology, Albert Ludwigs …, 2004

[45] P. Lall, D. Zhang, and J. Suhling, “High strain rate properties of SAC305 leadfree solder at high operating temperature after long-term storage,” in *2015 IEEE 65th Electronic Components and Technology Conference (ECTC)*, May 2015, pp. 640-651, doi: 10.1109/ECTC.2015.7159659

[46] K. Mysore, G. Subbarayan, V. Gupta, and Ron Zhang, “Constitutive and Aging Behavior of Sn3.0Ag0.5Cu Solder Alloy,” *IEEE Trans. Electron. Packag. Manuf.*, Vol. 32, No. 4, pp. 221-232, Oct. 2009, doi: 10.1109/TEPM.2009.2024119

[47] T. C. Lui, “Life-time prediction of viscoplastic lead-free solder: A new solder material, SACQ,” in *2017 IEEE International Workshop On Integrated Power Packaging (IWIPP)* Apr. 2017, pp. 1-4, doi: 10.1109/IWIPP.2017.7936754

[48] V. Ramachandran, K. C. Wu, and K. N. Chiang, "Overview Study of Solder Joint Reliablity due to Creep Deformation," *J. Mech.*, Vol. 34, No. 5, pp. 637-643, Oct. 2018, doi: 10.1017/jmech.2018.20

[49] Y. C. Chiou, Y. M. Jen, and S. H. Huang, "Finite element based fatigue life estimation of the solder joints with effect of intermetallic compound growth," *Microelectron. Reliab.*, Vol. 51, No. 12, pp. 2319-2329, 2011, doi: 10.1016/j.microrel.2011.06.025

[50] J. H. L. Pang, B. S. Xiong, and T. H. Low, "Low cycle fatigue models for lead-free solders," *Thin Solid Films*, Vol. 462-463, no. SPEC. ISS., pp. 408-412, 2004, doi: 10.1016/j.tsf.2004.05.037

[51] R. Al Athamneh, D. B. Hani, H. Ali, and S. Hamasha, "Reliability modeling for aged SAC305 solder joints cycled in accelerated shear fatigue test," *Microelectron. Reliab.*, Vol. 104, Jan. 2020, doi: 10.1016/J.MICROREL.2019.113507