Proceedings of IAC in Budapest 2021

Budapest, Hungary

November 26 - 27, 2021

Global Education, Teaching and Learning (IAC-GETL)
Engineering, Transport, IT and Artificial Intelligence (IAC-ETITAI)
Management, Economics, Business and Marketing (IAC-MEBM)

ISBN 978-80-88203-24-7

CTU
CZECH TECHNICAL UNIVERSITY IN PRAGUE

CIAE
Czech Institute of Academic Education z.s.

Impressum

Title:

"Proceedings of IAC in Budapest 2021", ISBN 978-80-88203-24-7

November 2021 in Prague, *1st edition*

Publisher / Creator of CD / Copyright holder:

Czech Institute of Academic Education z.s.

Address of Publisher:

Vodnicka 309/20, 149 00 - Prague 4, Czech Republic

Email: info@conferences-scientific.cz

Web: www.conferences-scientific.cz

Technical editorial staff this CD:

Helena Kratochvílová

Radek Kratochvíl, Ph.D.

Název:

"Proceedings of IAC in Budapest 2021", ISBN 978-80-88203-24-7

Listopad 2021 v Praze, *První vydání*

Vydavatel / Tvůrce CD / Nositel autorských práv:

Czech Institute of Academic Education z.s.

Adresa vydavatele:

Vodnická 309/20, 149 00 - Praha 4, Česká Republika

Email: info@conferences-scientific.cz

Web: www.conferences-scientific.cz

Technická redakce tohoto vydání:

Mgr. Helena Kratochvílová, DiS.

Ing. Radek Kratochvíl, Ph.D.

Conference Scientific Committee / Reviewers

Assoc. Prof. PhDr. Maria Janesova - Slovakia
- *Czech Technical University in Prague*

Mgr. Helena Kratochvilova - Czech Republic
- *Czech Institute of Academic Education*

Piotr Jednaszewski, Ph.D. - Scotland
- *St Mary`s Academy Ltd*

Assist. Prof. Mario Konecki, Ph.D. - Croatia
- *University of Zagreb, Faculty of Organization and Informatics*

Assoc. Prof. Souvik Pal - India
- *Department of Computer Science and Engg. in Nalanda Institute of Technology*

Assoc. Prof. Dr. Ioan-Gheorghe Rotaru - Romania
- *'Timotheus' Brethren Theological Institute of Bucharest*

Assoc. Prof. Dr. Ömer Beyhan - Turkey
- *Konya NE.University, A.K. Education Faculty*

Assoc. Prof. Anjali - India
- *Department of Computer Science & Engg, GNIT, Mullana*

Assoc. Prof. Jitender Grover - India
- *Department of Computer Science & Engg., M.M. University*

Radek Kratochvil, Ph.D. - Czech Republic
- *Czech Institute of Academic Education, Czech Technical University in Prague*

Assist. Prof. Dr. Mohamad H. Atyeh - Kuwait
- *School of Business, Australian College of Kuwait*

Assoc. Prof. Dr. Ramazan Sak - Turkey
- *Yüzüncü Yıl University, Faculty of Education*

Assoc. Prof. Dr. İkbal Tuba Şahin-Sak - Turkey
- *Yüzüncü Yıl University, Faculty of Education*

Assist. Prof. Dr. Ayşegül Derman - Turkey
- *Konya NE.University, A.K. Education Faculty*

Assist. Prof. Dr. Serdar Derman - Turkey
- *Konya NE.University, A.K. Education Faculty*

Assoc. Prof. Dr. Fethi Kayalar - Turkey
- *Erzincan University, Faculty of Education*

Prof. Dr. Thomas Rachfall - Germany
- *Hochschule Merseburg*

Dr. Dirk Förster-Trallo - Germany
- *Hochschule für Technik und Wirtschaft Berlin*

Editors and Technical publishing board:

Mgr. Helena Kratochvilova - Czech Republic
Radek Kratochvil, Ph.D. - Czech Republic

Organizing Committee:

Mgr. Helena Kratochvilova - Czech Republic
Assoc. Prof. PhDr. Maria Janesova - Slovakia
Radek Kratochvil, Ph.D. - Czech Republic

Table of Contents

Global Education, Teaching and Learning (IAC-GETL)

Role of Authentic Texts in Second Language Lexical Acquisition and Corpora-Based Approach to Language Teaching
Martina LIPKOVÁ **8 - 19**

The Kindergarten-Teachers' Leadership and Management during The Corona Pandemic (COVID 19): From Uncertainty to Caring
Aviva AVIDAN, Yonit NISSIM **20 - 22**

Importance of Enriched Virtual Learning Model in Hybrid Teaching Application
Fethi KAYALAR **23 - 27**

Engineering, Transport, IT and Artificial Intelligence (IAC-ETITAI)

Identification of Deceptive Opinion Spam Based on Deep Neural Networks
Aytuğ ONAN **29 - 36**

Deep Learning Based Identification of Offensive and Hateful Language on Social Media
Aytuğ ONAN **37 - 43**

Factory Environment Safety via Edge Computing Technology
Yang-Han LEE, Cheng-Wen CHEN, Yi-Lun CHEN, Chen-Shiuan WANG **44 - 54**

Correlation between Overall Equipment Effectiveness (OEE) and Failure Mode and Effect Analysis (FMEA) at The Semi-Automatic Assembly Lines
Péter DOBRA, János JÓSVAI **55 - 63**

A Data-driven Solution for The Dynamic Capacitated Arc Routing Problem
Zsuzsanna NAGY, Ágnes WERNER-STARK, Tibor DULAI **64 - 83**

Management, Economics, Business and Marketing (IAC-MEBM)

Global Energy Crisis: Impact on The Global Economy
Peterson K.OZILI, Ercan OZEN **85 - 89**

Impact of Internal Audit on Financial Performance: Commercial Banks
Arjeta HALLUNOVI **90 - 94**

Challenges, Opportunities for The Development of New Banking Products in Pandemics
Egzona HASANI, Besim BEQAJ **95 - 99**

Global Education, Teaching and Learning
(IAC-GETL 2021)

Role of Authentic Texts in Second Language Lexical Acquisition and Corpora-Based Approach to Language Teaching

Martina LIPKOVÁ [a]1

[a] *Slovak University of Technology in Bratislava, Faculty of Mechanical Engineering, Department of Languages and Social Sciences, Námestie slobody 17, 812 31 Bratislava, Slovak Republic, martina.lipkova@stuba.sk*

Abstract

The second language (L2) and specialist lexical acquisition are significant for university students' ability to participate in international exchange study programs, research cooperation, communicating around the topics concerning their study area.
This research study deals with the use of corpora and authentic texts as a major source of study materials for the development of communicative competence in English for Mechanical Engineering. Accurateness of particular specialty technical lexis and phrases used by the second-year university students with B2 proficiency level, in comparison with authentic texts compiled in corpora, has been analysed with the aim to present some strategies for corpora-based teaching of English for specific purposes (ESP). The research has been also realized in reaction to the results of university students' needs survey on the importance of studying ESP from their future career perspective, carried out by the Department of Foreign Languages and Social Sciences at the Faculty of Mechanical Engineering STU in Bratislava.
In the research, with the emphasis on authenticity, corpora-based approach to lexical acquisition was applied. Language analysis is based on the use of authentic texts and learner essays corpora on automotive industry. Connection between patterns and meaning enables students to understand a particular technical term as part of a phrase rather than in isolation that leads to better understanding of authentic materials. From this perspective, the lexical approach lies in teaching collocations, i.e. how lexical items co-occur and what grammar they tend to be associated with. In addition, knowledge of collocations supports effective sentence producing and enhances L2 proficiency development in students.

Keywords: authentic texts, collocations, corpora, ESP, lexical acquisition

1. INTRODUCTION

International mobility, cooperation and communication in various areas including education, culture, sports, industry, commerce, etc., have been realized in different forms with emphasis on intensified physical mobility of students, academic staff, researchers, etc. In higher education, international cross-border transfer of knowledge and intercultural competence contribute to overall universities' reputation along with offered study programs based on the latest research outcomes, strategies of education with focus on the students' competence to act in international environments.
Foreign language training is considered one of the major assumptions for international communication strategies development in higher education. International students' mobility is the indicator most referred to from the internationality university ranking perspective. Thus, foreign or L2 specialty lexis acquisition in students contributes to accomplishment of the objectives.

1.1. Authenticity in foreign language learning and teaching

Authentic situations, also referred to as genuine, real or natural, provide significant input for the foreign language acquisition with various authentic texts in different forms: auditory, print, visual and those that refer to realia.

* Corresponding author.

Definitions of authenticity prove how differently the concept of authenticity is perceived. It is mostly defined as the language produced by native speakers in a particular language community of real audience (Galloway 1992; Morrow 1977; Porter and Roberts 1981) or as the "spoken or written language data that has been produced in the course of genuine communication, and not specifically written for the purposes of language teaching" (Nuan 1999).

From the teacher's perspective, learning aims must be stated, i.e. to respond the questions 'What are we trying to achieve with authentic texts and further teaching materials based on the content and discourse they deliver? What is the language to be learnt so that a learner is able to reproduce the information the authentic text conveys? How can be the learner's output achieved by the use of authentic texts assessed? Which learning strategies contribute to lexical acquisition to enhance the learners' communicative competence? Does it enhance learners' motivation, higher-order thinking skills or greater learner autonomy? Can be the information communication technologies (ICT) applied as the means of lexical acquisition?

Foreign languages are the tools of communication through which an interaction and communication between scientific and technological communities is realized. Hereby, professionals from different scientific and industrial environments are expected to master high level of foreign language communicative competence. Technical universities play an important role in providing foreign language courses for students, in our case the future mechanical engineers, from the scientific cognition and technological knowledge perspective. Learning English for specific purposes (ESP) is regarded a complex psycholinguistic activity for students comprising specific aspects of academic reading – understanding of specific authentic texts, grammar structures and their functions, syntax and semantics.
English for mechanical engineering courses at STU in Bratislava provide the students with the input of language used in various technical areas, e.g. aviation, automotive industry, materials and their properties, mechanisms, environmental technologies, robotics and artificial intelligence, biotechnology, etc., and reflect the large scale of subject-based contents. Particular studying programs with specific subjects convey different discourse characteristics in English authentic texts that include very specific terminology to identify, e.g. the names of machine components designed by engineers, how the machines work, processes applied in innovative technologies and the activities starting from an idea to the marketplace. Specialty vocabulary of definitions, hypothesis, trends and events expressed by charts, mathematical symbols, equations or geometry specialty terms, etc., are learnt by students combining both the subject-based content, lexical acquisition with focus on communicative and task-based approach to promote motivation for learning foreign language in students. Thus, the use of authentic texts is advocated by many ESP teachers for their diversity and a wider range of vocabulary as the attempt to complete the gap between textbook and authentic discourse. Furthermore, language is a dynamic process in which technological innovations, new concepts, neologisms, updated information play an important role in ESP teaching, as glossaries and teaching materials should be adapted to these new conditions to keep pace with students.

The choice of authentic texts should provide the students with contexts to stimulate their interest in learning foreign languages. Pragmatic gains of foreign language proficiency support motivation, self-confidence, personal development and responsibility in students, thus enhancing their chances in the labor market.

2. ESP in tertiary education

2.1. Survey into the importance of learning foreign languages

Data about the graduates' level of foreign language proficiency and the measure to which they use foreign languages at their workplaces were obtained in a survey carried out by the Department of Foreign Languages and Social Sciences at the Faculty of Mechanical Engineering of the Slovak University of Technology in Bratislava, within which 20 companies in the tech industry employing graduates from different universities were surveyed. 45% of the companies confirmed that language skills mastered by the graduates they employ are sufficient, 25% of the overall number of the companies show neutral attitude, and 20% of them confirmed an insufficient level of foreign language proficiency. 45% of the companies confirmed that their employees use English as a foreign language every day, 30% partially, and 25% not every day. 80% of employers strongly confirmed the importance of communicative competence in English in their employees. Based on the facts mentioned above, the fusion of academic subject matter and specialist language communication competence is thought of as the supreme importance in education.

Data about the students' attitude toward foreign language learning were obtained in a survey carried out by the Department of Foreign Languages and Social Sciences, within which 85 second-year students were surveyed.
How students rank the importance of learning foreign languages at the university is displayed in Figure 1. Based on the results, 46% of the surveyed students confirmed that it is important and 36% ranked it very important. 14% of the students reported that learning languages is necessary to be able to perform the job in their future career.

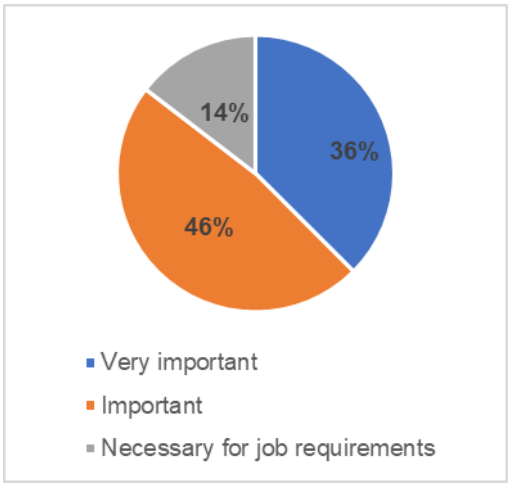

Fig. 1. The importance of learning foreign languages at the university

The rate and type of motivation for learning languages is shown in Figure 2. Based on the survey, the percentage of students that learn languages to enhance communicative competence and find a job in a foreign company is 67%, as students responded to either of the options offered in the questionnaire. The percentage of the students who wish to work in an international company located in the Slovak republic is 48%.

Fig. 2. Reasons to learn languages by the university students

The students were also asked to state their views of the importance to achieve a high level of foreign language proficiency from their future career perspective on the scale: very important, important and neutral. As shown in Figure 3, the largest percentage is reported as important (42%) and very important (39%), compared to 13% of those who were neutral in their responds.

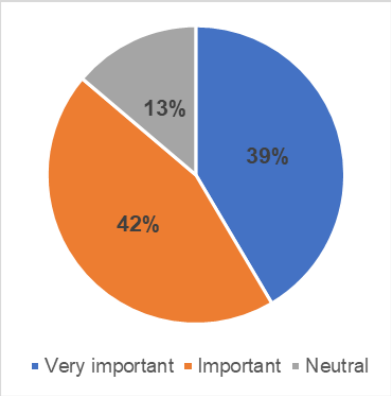

Fig. 3. The importance of high foreign language proficiency for students' future career

Despite the survey outcomes displayed in Figure 3 confirming that the students regard learning foreign languages associated with the high proficiency of foreign language communication important and even very important, only 45% of the surveyed companies are satisfied with sufficient language communication skills in their employees. Due to findings gained from the survey, a different and innovative approach to teaching and learning foreign languages, including specialty lexical acquisition, should be considered. For this purpose, corpora-based approach has been applied to design activities aimed at proper use of specialty technical terminology and discourse development.

2.2. ESP at technical universities

Learning and teaching languages at technical universities covers the activities with the aim to help students communicate in their studying-area-based communication and to use it also later in their professional career. Both the context and language structures are important in learning and teaching ESP. A wide range of subjects is included, therefore a need analyses has been worked out to develop language skills upon the future job requirements.

Academic writing is trained by students in the form of written projects based either on knowledge achieved through practical classes in laboratories, or studios when designing machines or their components in combination with studying authentic texts in English.

The incorporation of authentic technical texts adapted for teaching purposes provides a good opportunity for understanding the specific meaning of terms within a discourse (Table 1). Thus, authentic technical texts in a foreign language characterized by a high concentration of technical words with their explicit meaning, diagrams, charts, etc. represent valuable teaching materials. Working with text may be realized in the form of team-based activities working on a common project that requires looking up and sorting information related to the task. Interpretation and presentation of the students' findings with their attitude to a specific problem develop both critical thinking and motivation in students. They become more enthusiastic about learning academic foreign language by discussing an issue and listening to the attempt to understand the utterances of other students (Lipková 2020).

Table 1. Specific field-of-study-based terminology within the discourse

Terms used to describe, for example, properties of materials; shear loading; torsion loading in mechanical engineering:	
shape	Materials change their *shape* and size.
properties	Engineers have to take into account mechanical *properties* of materials.
load	Some materials resist deformation under the application of *loads.*
deformation	*Deformation* refers to changes in the shape or size of an object.
distortion	There's a small *distortion* in the shape of this object.
wear &tear	Wear & tear is a damage that occurs naturally when an object is used; it is exposed to external forces.
rotation	The *rotation* of the cross section at the free end of the shaft is called the angle of twist.
torque	The *torque* refers to the twisting moment.
Some terms and their definitions within the discourse:	
high elastic modulus	A *high elastic modulus* is typical for materials that are hard to deform; in other words, materials that require a high load to achieve a significant strain; for example, a steel band.
low elastic modulus	A *low elastic modulus* is typical for materials that are easily deformed under a load; for example, a rubber band.
stress	The *stress* applied to a material is the force per unit area applied to the material. The maximum stress a material can stand before it breaks is called the breaking stress.
strain	*Strain* is a measure of material deformation in response to an applied force (or stress).
Hooke's law	*Stress* is directly proportional to strain.

3. Corpora-based approach in teaching ESP

To understand and exploit effectively the language analyses options offered by corpora, the definition of corpus and corpus linguistics might help teachers to see the potential of corpora-based approach in educational linguistics, either applied directly in the form of in-class activities, or a valuable source of data-driven outcomes through the queries given by teaching objectives resulting from learners' needs. A *corpus* is defined as "a large and principled collection of natural texts created systematically for particular purposes" (Anderson and Corbett 2017; Biber et al. 1998). It is also defined as "a collection of linguistic data (usually contained in a computer database), used for research, scholarship and teaching. Also called a text corpus" (Nordquist 2020). Anderson and Corbett (2017) define *corpus linguistics* as "a set of methods for the analyses of language based on the evidence provided by corpora of texts". According to O'Keeffe and McCarthy (2012), it "demonstrates and supports linguistic statements and assumptions". Based on the W3-Corpora Project (University of Essex 1998), *corpus linguistics* is defined "as the study of linguistic phenomena through large collections of machine-readable texts".

The results for queries can be obtained in different forms. For example, Word Sketches (Sketch Engine) give a summary of word usage, i.e. which other words commonly co-occur with the keyword, furthermore in what grammatical relation they are naturally used. For example, the verb "*drive"* can be used with prepositions: *„Hybrid electric cars are* ***driven by*** *the engine to generate electricity"; "Working vehicles* ***drive at*** *low speed"*. Another concordance may be found in a sentence, where *subject + "drive"* occurs: *"The propulsion engines* ***drive*** *the propellers via toothed belts".* Another results may be found on statistically frequent, or infrequent, words in different registers of texts, and groups of words that appear frequently together can be found in cluster lists. For example, in mechanical engineering, the following keyword list (Figure 4) of statistically frequent nouns, has been found in a corpus containing authentic texts on automotive industry that was created for educational purposes at the Faculty of Mechanical Engineering:

noun (51,472 items | 1,108,855 total frequency)

	Lemma	Frequency		Lemma	Frequency		Lemma	Frequency		Lemma	Frequency
1	vehicle	20,842	14	component	4,806	27	truck	3,294	40	year	2,544
2	system	15,845	15	part	4,348	28	type	3,045	41	manufacturer	2,520
3	brake	10,153	16	time	4,346	29	process	3,029	42	use	2,497
4	design	8,652	17	speed	4,186	30	application	2,979	43	load	2,485
5	car	8,104	18	chassis	4,070	31	air	2,888	44	mm.	2,484
6	engine	7,397	19	[number]	3,892	32	company	2,871	45	safety	2,413
7	power	6,739	20	wheel	3,886	33	manufacturing	2,838	46	pressure	2,400
8	control	6,030	21	model	3,854	34	driver	2,828	47	information	2,352
9	motor	5,705	22	material	3,729	35	performance	2,709	48	level	2,345
10	battery	5,386	23	energy	3,593	36	drive	2,678	49	datum	2,342
11	figure	5,309	24	market	3,581	37	cost	2,609	50	service	2,319
12	fuel	5,136	25	product	3,478	38	industry	2,601			
13	technology	4,838	26	sensor	3,444	39	development	2,584			

Fig. 4. Sketch Engine Wordlist – Mechanical Engineering Automotive Industry Corpus

For educational purposes, the keyword list created in the Sketch Engine tool displayed in Figure 4 can be compared with the keyword list of a learner corpus (Figure 5) in which the students' final semestral essays on automotive industry have been compiled:

noun (793 items | 3,071 total frequency)

	Lemma	Frequency		Lemma	Frequency		Lemma	Frequency		Lemma	Frequency
1	car	220	14	[url]	25	27	example	18	40	way	11
2	engine	138	15	transmission	25	28	world	17	41	charge	11
3	battery	98	16	electricity	25	29	advantage	17	42	manufacturer	11
4	vehicle	91	17	range	24	30	road	16	43	air	11
5	combustion	55	18	figure	23	31	environment	15	44	cylinder	11
6	emission	41	19	part	22	32	pollution	15	45	time	10
7	technology	40	20	future	21	33	pack	15	46	traction	10
8	energy	35	21	work	21	34	year	14	47	audi	10
9	power	32	22	people	20	35	problem	14	48	life	10
10	type	30	23	station	20	36	electric	13	49	petrol	10
11	production	27	24	ev	19	37	transport	12	50	capacity	9
12	system	26	25	april	19	38	tesla	12			
13	fuel	26	26	motor	19	39	disadvantage	12			

Fig. 5. Sketch Engine Wordlist – Mechanical Engineering Learner Corpus

As a result of the analyses, in Figures 4 and 5, the "DNA" characteristic for the texts in this particular discourse in both the web and learner texts on automotive industry can be seen in the two corpora mentioned above. Thus, we can see that the most frequent word - *car* - was preferred by students in their writing texts compared to *vehicle* used the most frequently in the web texts. Another significant finding refers to the frequency of the noun *engine*, i.e. how often it appeared in the corpora. In the learner corpus, it occurs on the second position (the absolute frequency of 138 occurrences out of 3,071 total frequency), but in the web texts corpus, the same noun is displayed on the sixth position (with the absolute frequency of 7397 out of 1,108,855 total frequency). So, what is the reason for such a difference in the use of the same word in a common field of study, that is, automotive industry? Does it result from specific topics preferred by the students with regards to different types of engines? Are the students interested particularly in the engine systems, or their design, production, etc.? To answer the questions, further types of analyses can be done, including cluster lists frequencies, selected keyword collocations, consistency and dispersion of words that reoccur consistently in lots of texts of a given genre. In addition, examples of the use in context can be compared in the corpora by the means of concordance corpus search tool. This is often used to teach learners how to process language data inductively, i.e. make generalization based on observing examples, thereby discovering the rules.

4. Methodology

For the purpose of this research, three types of corpora have been used for further lexico-grammar analyses and designing ESP teaching tips for the technical English courses at the Faculty of Mechanical Engineering. English for mechanical engineering includes specialty terms and structures characteristic for the discourse. A list of the most frequent words, single or multi-word terms, word sketch collocations and word combinations, concordance examples used in context and parallel concordance tool have been used to compare the word behaviour in three different corpora: British Academic Written English Corpus (BAWE), Mechanical Engineering – Automotive Industry Corpus, Mechanical Engineering Learner Corpus.

The BAWE is the ready-to-use preloaded corpus included in Sketch Engine, and the next two: Mechanical Engineering - Automotive Industry Corpus and Mechanical Engineering Learner Corpus have been created for educational purposes, i.e. English for mechanical engineers. The last of the previously mentioned corpora helped find out the differences in the way of how the selected words were used by the students themselves in writing their technical essays compared with the word occurrence in authentic texts. In addition, the BAWE corpus key single and multi-terms, as well as other word analyses, were applied to compare the word query outcomes with those in the Mechanical Engineering Learner Corpus included. The BAWE corpus contains 2,761 pieces of assessed student writing (Anderson and Corbett 2017), including various academic fields of study, thereby chosen to accomplish the research objectives.

4.1. Sample description

Second-year students with B2 (upper-intermediate) communication proficiency level of English were assigned to write a technical essay on the topic with regard to their field of study at the Faculty of Mechanical Engineering. As a result of the topic ideas for technical academic writing preferred by the students, those focusing on automotive industry proved as attractive to students, thus providing a good source of the students' texts to create a learner corpus dealing with various aspects of car construction, components, systems, etc.

4.2. Corpora characteristics

In Figures 6, 7 and 8 (Sketch Engine - Corpus Info), basic descriptions of the newly created, as well as preloaded BAWE corpora are shown. The number of words and frequencies result from different options provided by the corpora. The learner corpora depends on limits given by a specific discipline area – automotive industry - for technical writing in learner corpora. The students writing was also selected with regards to designing some teaching tips, where the authentic texts from websites and the texts written by students were collected for further lexico-grammar analyses. In terms of language teaching, real and authentic context examples compared with the texts created by students may increase motivation in students through searching for correct and natural use of specialist words and grammar structures in a particular real-life context.

COUNTS

Tokens	11,694
Words	10,055
Sentences	547
Documents	23

Fig. 6. Sketch Engine corpus info – Mechanical Engineering Learner Corpora

COUNTS

Tokens	8,336,262
Words	6,968,089
Sentences	293,113
Paragraphs	127,401
Documents	2,761

Fig. 7. Sketch Engine corpus info - British Academic Written English Corpus (BAWE)

COUNTS

Tokens	3,347,245
Words	2,717,629
Sentences	143,909
Paragraphs	47,061
Documents	1,268

Fig. 8. Sketch Engine corpus info – Mechanical Engineering – Automotive Industry

5. Corpora keywords analyses

The Mechanical Engineering Learner Corpora is a collection of 23 texts written by students and selected from the texts written by students including various topics concerning different fields of study in mechanical engineering. The 23 selected students written texts on automotive industry make the main body of the learner corpus in this research to represent a sample of a language used by university students in academic writing. The learner corpus was created with a linguistic purpose to exploit the options of lexical acquisition in students based on comparison with the corpus compiled from the web material, also focusing on automotive industry. The BAWE corpus, though containing students written texts on various topics, is another source used to compare and enhance academic writing in university students. To characterise the main characteristics of the corpora involved, the following comparison of 10 keywords related to automotive industry have been extracted (Table 2).

Table 2. First 10 keywords (different parts of speech) in the Sketch Engine corpora with regards to automotive industry

	Mechanical Engineering Learner Corpus (compiled from students' essays)	Mechanical Engineering – Automotive Industry Corpus (compiled from web texts)
1	combustion	heavy-duty
2	wankel	brake
3	škoda	chassis
4	lithium-ion	chassi
5	ev	hev
6	woollacott	autonomous
7	rs3	hybrid
8	electric	ecu
9	charging	powertrain
10	suvs	automotive

The frequency of keywords in corpora is statistically significant, thus providing words which are typical for the used corpora. The keywords in the learner corpora (Table 2) seem to be the words being used the most frequently by the students when writing the texts and reveal the most frequent topics for writing. Different keywords are extracted from the corpora compiled from web texts, pointing out to other most frequently used terms in the texts on the same engineering area, i.e. automotive industry. These might be thought of the next important list of words to be studied by students in particular contexts.

For example, the frequency of the word *combustion (N = noun)* is represented by 55 occurrences, and is on the highest learner corpus position in the keywords list containing different parts of speech, as shown above (Table 2). The word *car (N)* occurred 220 times, as shown in the highest frequency words (nouns) in Table 3. Considering the Mechanical Engineering Learner corpus, it is interesting to find out how the most frequent words relate, thus the frequency list of words is also essential.

Table 3. The highest frequency words (nouns) in the learner corpus

	Mechanical Engineering Learner Corpus	Occurrence
1	Car	220
2	Engine	138
3	Battery	98
4	Vehicle	91
5	combustion	55

If we search for the key multi-word terms, the most frequent nouns from Tables 2 and 3 occur, as follows:

Table 4. The key multi-word terms in the learner corpus

	Mechanical Engineering Learner Corpus
1	Electric car
2	Combustion engine
3	Internal combustion
4	Internal combustion engine
5	Traction battery
6	Battery pack
7	Engine technology
8	Electric vehicle
9	Traction battery pack

10	Electric traction motor

Having extracted the lists of keywords, key multi-word terms and the most frequent words, as the first step in searching for the main characteristics of corpora, the context in which particular specialist vocabulary occurs is a useful tool for better understanding of its natural use in the language. The example sentences for the following terms are displayed in Table 5.

Table 5. Sample sentences retrieved from the Sketch Engine Concordance tool

Combustion (N.) in sample concordance sentences extracted from the Mechanical Engineering Learner Corpus (compiled from the students' written essays)
Combustion engines are supposed to be replaced by electric motors.
As everyone knows, without ***combustion engines***, everything would be different, …
… aimed at comparison of the advantages and disadvantages of an internal ***combustion engine***.
The internal ***combustion engine*** is one of the most complex but also the most used types of engine.
During the ***combustion of the mixture***, a number of harmful substances are released.

As displayed in Table 5, the sample sentences created by students contain mostly the collocations of the noun *engine* modified by *combustion*, i.e. *combustion engine*, and in one case *combustion* was followed by the preposition *of (the combustion of the mixture).* As a result, teaching materials focusing on lexical acquisition concerning automotive industry, as a part of the mechanical engineering discipline, should provide students with more naturally used options, thus enhance the communication proficiency. For this purpose, more examples of word combinations might be included. For example, the word combinations including *combustion*, extracted from the corpus compiled from the web texts, are shown in Table 6.

Table 6. Web texts corpus collocations

Combustion collocations from the Mechanical Engineering – Automotive Industry Corpus, compiled from web texts)	
Nouns modified by *combustion*	*The combustion chamber; the combustion process; combustion temperature; combustion stroke; combustion quality, etc.* Sample concordance sentences: *A piston ring seals the **combustion chamber** through inherent and applied pressure.* *Explain the **combustion process** in a diesel engine.*
Modifiers of *combustion*	*The incomplete combustion; stable combustion; efficient combustion; rapid combustion; fuel combustion; etc.* Sample concordance sentences: *Carbon monoxide results from the **incomplete combustion** of hydrocarbons due to a lack of oxygen.* *This results in more **efficient combustion**, a 2–5% fuel economy and proportionate reduction of CO2 emissions.*
Verbs with *combustion* as an object	*Ensure complete combustion; accelerate the combustion; undergo combustion; improve combustion etc.* Sample concordance sentences: *This helps to **ensure** balanced **combustion.*** *Conversion of free energy in the fuel into electric energy, without **undergoing combustion.***

Further approach to analyses might be applied through exploiting the way how synonyms, e.g. *engine* and *motor*, are used in real situations, as they occurred in the list of key multi-word terms extracted from the corpus compiled from authentic website texts.

Table 7. Comparison of the Sketch Engine collocations of two words: *engine* and *motor*

Modifiers of the words		
	engine	*motor*
combustion	373	0
internal	244	0
diesel	154	0

gasoline	122	0
petrol	87	0
IC	81	0
electric	12	856
stepper	0	67
magnet	0	84
traction	0	137
induction	0	131
DC	0	200

As shown in Table 7, the collocations *combustion engine (373 occurrences), internal engine (244 occurrences), diesel engine (154 occurrences), etc.,* but *stepper motor (67 occurrences), magnet motor (84 occurrences), traction motor (137 occurrences), etc.,* are used in authentic texts on automotive industry. However, *electric engine* collocation is represented by 12 occurrences, and *electric motor* by 856 occurrences, what helps to decide for the collocation that is more natural in a particular field of study. The Sketch Engine tool also provides teachers/learners with further functions, e.g. collocation concordances (how they are used in sentences), collocation word sketch (how they are used in combination with other words, e.g. *traditional/ internal/ external combustion engine*), and Thesaurus tool used for extracting synonyms and similar words.

When using parallel concordance, translation of a particular word or terms can be seen in another language upon the choice from the given options. DGT-Slovak corpus based on the DGT-Translation Memory database (Steinberger et al. 2013) was used to create parallel concordance. In this research, e. g. the sentences including *electric motor* (Table 8) and *electric engine* (Table 9) in English were compared to their translated equivalents in Slovak language. The Slovak term *'motor'* (and its particular forms) is used in the Slovak translated equivalents for the English collocations: *electric engines* and *electric motors.* Based on the examples below, students are encouraged to find out a general rule on the use of the selected collocations in English depending on the context.

Table 8. Parallel English-Slovak (DGT corpus) Sketch Engine concordance for *electric motor*

Fans (excluding table, floor, wall, ceiling or roof fans with a self-contained ***electric motor*** of an output ≤ 125W, axial fans, centrifugal fans) …	Ventilátory (s výnimkou stolových, podlahových, stenových, okenných, stropných alebo strešných ventilátorov so samostatným *elektrickým motorom* s výkonom ≤ 125W, osových, odstredivých) ...
Domestic vacuum cleaners with a self-contained ***electric motor*** for a voltage ≥ 110V	Domáce vysávače prachu so samostatným *elektrickým motorom* pre napätie ≥ 110V
Table, floor, wall, window or roof fans, with a self-contained ***electric motor*** of an output ≤ 125W	Stolové, podlahové, stenové, okenné, stropné alebo strešné ventilátory s vlastným *elektrickým motorom* s výstupom ≤ 125W

Table 9. Parallel English-Slovak (DGT Corpus) Sketch Engine concordance for *electric engine*

New and innovative solutions are necessary, based on ***electric engines*** and batteries, hydrogen and fuel cells, gas-powered engines, advanced architectures and technologies in engines or hybrid propulsion.	Potrebné sú nové a inovačné riešenia založené na *elektrických motoroch* a batériách, vodíku a palivových článkoch, plynových motoroch, pokročilých architektúrach a technológiách v motoroch alebo hybridnom pohone.
'propulsion' means a combustion engine, an ***electric engine***, any hybrid application or a combination of those engine types or any other engine type; …	„pohon" je spaľovací motor, *elektrický motor*, akákoľvek hybridná aplikácia alebo kombinácia uvedených typov motorov alebo ktorýchkoľvek iných typov motorov; …
'maximum continuous rated power' means the maximum thirty minutes power at the output shaft of an ***electric engine*** as set out in UNECE regulation No 85; …	„maximálny trvalý menovitý výkon" je maximálny tridsaťminútový výkon na výstupnom hriadeli *elektrického motora* podľa predpisu EHK OSN č. 85; …
Propelled with an ***electric engine***; …	Poháňané *elektrickým motorom*; …

6. Corpora-based classroom activities

Well-designed activities leading to enhancement of learner communication skills in a foreign language, as well as lexical acquisition in students of mechanical engineering, are essential for building autonomous learning in order to achieve a high level of the specialist and subject-based knowledge of terminology and its use in a real life. In this regard, data-driven analyses based on ready-to-use or user-created corpora provide an important source of authentic texts intended for various purposes.

Teaching materials for classroom activities can be designed with the help of corpora, the use of which brings a shorter time needed for investigating and preparing correct language within the discourse.

With regards to authentic text corpora, multiple-choice, gap-filling or fill-in-the blanks exercises, simple substitution tables, word formation, translation exercises, etc., can be created in a shorter time. Advanced users of corpora can get

the benefit out of corpora by access to a high number of genuine texts in the target language. Corpus Query System (CQS) enables teachers, students, translators and other users to work with large amounts of texts according to their needs and objectives.

From this perspective, computational linguistics and corpora-based teaching of L2 might be thought of an effective solution for teachers in their effort of designing study materials that meet teaching/learning objectives and learners' needs. This can be used to facilitate designing ESP lessons for professionals, as well as content and language integrated learning (CLIL) which has been regarded as an effective symbiotic fusion of studying particular academic subject-matter content (e.g. mathematics; physics; history; geography; etc.) through a foreign language.
Collecting ideas around a particular topic may be also achieved with the use of corpora applied in mind-mapping strategy when clustering ideas and defining relationships help to develop students' motivation and speaking skills while achieving both content and language objectives (Buzan 1996).

Conclusion

The use of corpora is becoming increasingly popular in language teaching and learning for containing large collections of texts providing various types of approaches to language analyses. Corpora, with authentic data involved, provides a valuable source for teachers and textbook authors to design syllabuses, teaching materials and tests based on data-driven information. Language structures and phrases typical for various text registers, or types of discourse, allow to answer great number of research questions and hypothesis, or verify our intuition about the language. Corpora brings an opportunity for teaching/learning ESP that is most relevant for the L2 learners due to advantages offered by corpora in searching for specialty lexis, grammar structures, the most frequent word collocations, and further language analyses upon the corpora queries. From the mechanical engineering perspective, the ESP teachers attempt to increase motivation in students to learn the structures needed for very specific, occupational and professional purposes.

Acknowledgements

The author gratefully acknowledges the contribution of the Cultural and Educational Grant Agency (KEGA) MŠVVaŠ SR under the grant 034STU-4/2019

References

[1] Anderson, W. & Corbett, J. (2017). *Exploring Engish with Online Corpora*. Second Edition. PALGRAVE. Macmillan Publishers Limited. ISBN 978-1-137-43809-6
[2] Biber, D., B., Conrad, S. & Reppen, R. (1998). *Corpus Linquistics*. Cambridge University Press. ISNB 0 521 49957 7
[3] Buzan, T. (1996). *The Mind Map Book. How to use radiant thinking to maximize your brain's untapped potential*. New York: Plume. ISBN 9780452273221
[4] Buzan, T. (1996). *The Mind Map Book*. *How to use radiant thinking to maximize your brain's untapped potential*. New York: Plume. ISBN 9780452273221
[5] Corpus Linguistics. W3-Corpora Project at the University of Essex. (1998). [online]. Accessed on: 19 July 2021. Retrieved from https://www1.essex.ac.uk/linguistics/external/clmt/w3c/
[6] Galloway, V. (1992). Toward a cultural reading of authentic texts. *Languages for a multicultural world in transition*, pp. 87-121. Lincolnwood. National Textbook Co.
[7] Gilmore, A. (2007). Authenic materials and authenticity in foreign language learning. *Language Teaching.* Volume 40. Issue 2. April 2007. pp 97-118. Cambridge University Press. doi:10.1017/S0261444807004144.
https://www.researchgate.net/publication/231910134_Authentic_materials_and_authenticity_in_foreign_language_learning
[8] Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., Suchomel. V. The Sketch Engine: ten years on. *Lexicography*, 1: 7-36, 2014. Retrieved from: https://www.sketchengine.eu/wpcontent/uploads/The_Sketch_Engine_2004.pdf
[9] Lipková, M. (2020). *Mechanical Engineering Lexis and CLIL Principles in Acquisition of Content-Based Lexical Competence in English*. DOI: 10.18355/XL.2020.13.02.15 in XLinguae. European Scientific Language Journal. Volume 13. Issue 2. Pp. 176-188. ISSN 1337-8384. eISSN 2453-711x. In SCOPUS: 2-S2.0-85084454440 (2019: 0.318 - SJR, Q1 - SJR Best Q). Retrieved from: file:///C:/Users/CJ2/Documents/%C4%8Cl%C3%A1nky/XLinguae2_2020_15.pdf
[10] Morrow, K. (1977). Authentic texs and ESP. In S. Holden (ed.). *English for Specific Purposes.* London. Modern English Publications. pp. 13-17.
[11] Nordquist, Richard. (2020). *Definition and Examples of Corpora in Linguistics*. Accessed on 19 July 2021. Retrieved from https://www.thoughtco.com/what-is-corpus-language-1689806
[12] Nordquist, R. (2020). *Definition and Examples of Corpora in Linguistics*. Accessed on 19 July 2021. Retrieved from https://www.thoughtco.com/what-is-corpus-language-1689806
[13] Nuan, D. (1998). Second language teaching and learning. Boston: Heinle and Heinle Publishers. ISBN-13: 978-0838408384
[14] Porter, Don & Jon Roberts. (1981). Authentic listening activities. *ELT Journals.* 36/1. Pp. 139-141.
[15] Porter, D. & Roberts, J. 1981. Authentic listening activities. *ELT Journals.* 36/1. Pp. 139-141.

[16] Schubert, L. (2020). Computational Linguistics, *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition), Edward N. Zalta (ed.). Retrieved from https://plato.stanford.edu/archives/spr2020/entries/computational-linguistics/
[17] Scrivener, J. (2010). Teaching English Grammar. What to Teach and How to Teach it. Macmillan Books for Teachers. Macmillan Publishers Limited. ISBN 978-0-2307-2321-4
[18] Sketch Engine. British Academic Written English Corpus (BAWE). Retrieved from: https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fbawe2
[19] Sketch Engine. DGT-Slovak, Slovak part of DGT-Translation Memory consisting of 24 languages. Retrieved from: https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fdgt__slovak
[20] Sketch Engine. Learn How Language Works. Lexical Computing CZ s.r.o. Retrieved from: https://www.sketchengine.eu/
[21] Sketch Engine. Mechanical Engineering_ Automotive Industry. Retrieved from: https://app.sketchengine.eu/#dashboard?corpname=user%2Flipkova%2Fmechanical_engineering_automotive_industry
[22] Steinberger, R., Eisele, A., Klocek, S., Pilos, S., & Schlüter, P. (2013). DGT-TM: A freely available translation memory in 22 languages. *arXiv preprint arXiv:1309.5226*.
[23] Teichler, U. (2017). Internationalisation Trends in Higher Education and the Changing Role of International Student Mobility. *Journal of international Mobility*, 1(1), 177-216. https://doi.org/10.3917/jim.005.0179
[24] Thornbury, Scott. (2002). *How to Teach Vocabulary*. Pearson Education Limited, pp. 185. ISBN 0582 429668

The Kindergarten-Teachers' Leadership and Management during The Corona Pandemic (COVID 19): From Uncertainty to Caring

Aviva AVIDAN [a]1, Yonit NISSIM [b]

[a] *Faculty of Education, Gordon College of education, Haifa, Israel. avivaa@godon.ac.il*
[b] *Faculty of Education telhai College, Katzrin, Israel. yonitn@telhai.ac.il*

Abstract

This pioneering research aims to shed some light and examine in a reflective manner the points of view of 64 kindergarten-teachers in Israel, regarding their leadership during the era of uncertainty in the outburst of the Corona pandemic. The research aims to analyze the way they characterize their leadership in central issues related to their educational management in time of a crises as leaders in the early childhood education. Their coping strategies with the complex situation forced them to recruit resources, personal strength, and different practices from times of normality and adopt them to the new reality. This strategy of coping is in accordance with LaBoskey & Richert's approach (2002), who advocate reflection as a methodological tool enabling alternative perspectives for a redesign of existing paradigms.

Keywords: kindergarten teachers, educational management, adaptive leadership, early childhood education.

1. INTRODUCTION

Observing the current time provides researchers alongside with the challenges of leading and managing, with the opportunity for learning, for professional growth and for purification of the managerial image standing as the head of the educational institution in times of uncertainty and crisis. Leadership in general, and adaptive leadership, in specific, have become key proficiencies in dealing with the different processes and the uncertainty situations in the current world pandemic (Laeur et. al, 2021; Nissim & Simon, 2020 ;2021). The 64 kindergarten-teachers embraced the adaptive leadership style and expressed it mostly in various ways of caring for their staff and the children's well-being (Gilligan, 1982: Noddings, 2005).

Studying the working ways of these kindergarten-teachers may provide us with numerous layers of knowledge regarding the domains of education, leadership and management, teacher and kindergarten-teachers' training (Avidan, 2017). It might also provide us with a different perspective of their professional deeds as educational leaders at a unique era (Kaden,2020).

1.1. Research Methodology: Quantitative

Statistical tests were performed as required in the quantitative research protocol. The study was based on a validated questionnaire distributed to a convenience sample of 64 kindergarten-teachers in different parts of Israel.

* Corresponding author.

1.2. Main findings

The main findings indicate that the kindergarten teachers reported high leadership characteristics across all metrics examined (team relationship, leadership qualities, concern and caring, and coping with change) with the highest-scoring index being concern and caring. For kindergarten manger (4.87) kindergarten teachers (4.81) the caring strategy received the highest score compared to complementary kindergarten teachers (4.61) and significantly F (2,61) = 2.66, p> 0.05.

Table 1. Averages and standard deviations of all questionnaire statements (N=64)

Statements	average	Standard deviation
1. The kindergarten teacher enables freedom of choice and a wide activity span to her staff	4.34	**.65**
2. The kindergarten teacher expresses respect and consideration to the children, her staff and the parents	4.87	**.34**
3. The kindergarten teacher sets a mutual trust relationship between her and the staff	4.84	**.41**
4. The kindergarten teacher leads processes of teaching and learning	4.72	**.52**
5. The kindergarten teacher shares with the parents and her staff all the challenges and problems in the kindergarten	3.77	**.94**
6. The kindergarten teacher expresses a caring attitude to her staff and to the children in her everyday activities	4.80	**.41**
7. The kindergarten teacher leads processes of developing an utmost educational classroom climate	4.88	**.33**
8. The kindergarten teacher is a leading figure who guides the parents and her staff during uncertain times	4.75	**.51**
9. The kindergarten teacher is not afraid of the challenges during uncertain times	3.59	**.99**
10. The kindergarten teacher demonstrates leadership skills	4.53	**.73**
11. The kindergarten teacher functions well under pressure situations	4.38	**.70**
12. The kindergarten teacher sets a personal example to the children and her staff	4.80	**.44**
13. The kindergarten teacher has self-confidence	4.44	**.75**
14. The kindergarten teacher demonstrates flexibility and functions well during extreme situations	4.55	**.71**
15. The kindergarten teacher is willing to and capable of affecting others	4.48	**.62**
16. The change that the staff and the kindergarten teacher have undergone during the uncertain time of Covid 19 was meaningful and challenging	4.19	**.80**
17. To what extend was the change reflected in the way the kindergarten teacher led and managed her staff	3.80	**1.01**
18. To what extend was the change expressed in the use of technological means and teaching methods	4.34	**1.02**

Conclusions

The key findings suggest that kindergarten teachers perceive themselves as leaders. A very high, self-esteem was found. The reliability of the indices, as examined according to alpha-Kronbach (α), were found to be high, a figure

that characterizes a high degree of stability and consistency in the respondents' responses to the statements of each of them.

The more significant the role of the kindergarten teacher, and the more days she is in kindergarten, the more significant her leadership is in all the indicators examined. It was found that there are positive and significant positive relationships between the relationship index and the indices of concern and caring and leadership qualities, in such a way that as the principal perceives herself as demonstrating caring and caring skills, she creates an affinity for her leadership qualities and expression at the level of her relationship.

A positive and moderate intensity was also found between the measure of concern and caring and the measure of leadership qualities, in such a way that the more caring the principal is, the more leadership qualities she exhibits. Her self-perception as an educational leader and not just as a kindergarten principal is built and developed as she learns to demonstrate the same qualities required of a leader in a time of crisis and uncertainty.

References

[1] Avidan, A. (2017). The role and contribution of narrative interviews in educational research. American Journal of Educational Research, 5, (4), 419-427. http://article.scieducationalresearch.com/pdf/EDUCATION-5-4-10.pdf

[2] Gilligan, C. (1982). In a different voice: psychological theory and women's development. Cambridge: Harvard University Press.

[3] Kaden, U. (2020). COVID-19 School Closure-Related Changes to the Professional Life of a K–12 Teacher. Education Sciences, 10(6), 165-178. https://doi.org/10.3390/educsci10060165

[4] LaBoskey, V. K., & Richert, A. E. (2002). Identifying good student teaching placements: A programmatic perspective. Teacher Education Quarterly, 29(2), 7-34.

[5] Laur, C.V, Agarwal, P, Mukerji G, Goulbourne, E, Baranek, H, Pus, L, Bhatia, R.S, Martin D, Bhattacharyya, O (2021). Building health services in a rapidly changing landscape: lessons in adaptive leadership and pivots in a COVID-19 remote, monitoring program. J Med Internet Research, 23(1):e25507.

[6] Nissim, Y., & Simon, E. (2021). Flattening the hierarchy curve: Adaptive leadership during the Covid-19 pandemic- A case study in an academic teacher training college. Review of European Studies, 13(1), 103-118.

[7] Nissim, Y., & Simon, E. (2020). Agility in teacher training: Distance learning during the Covid-19 pandemic. International Education Studies, 13(12), 11-25.

Noddings, N. (2005). The challenge to care in schools: An alternative approach to education (2nd edition). New York, NY: Teachers College.

Importance of Enriched Virtual Learning Model in Hybrid Teaching Application

Fethi KAYALAR [a]1

[a] *Assoc. Prof. Dr., (Erzincan B. Y. University), Faculty of Education,, Erzincan, TURKEY*
fkayalar@erzincan.edu.tr

Abstract

The Enriched Virtual Learning Model, which is included in the Hybrid Teaching application, which has been widely used in distance education due to the Pandemic, is very simple to implement compared to other models. The basis of student learning is online and requires the student to attend school only on certain days. Thanks to technology, learning takes place everywhere, at all times, and sometimes at varying speeds. Face-to-face practice, which forms part of this model, generally serves two main purposes; the first is to enrich students' learning experiences with group-based work or teacher guidance, and the second is to hold students accountable through regular face-to-face checkups and to meet with their teachers and advisors. This model is highly functional by nature. In our study, we aimed to determine how effective and efficient the Enriched Virtual Learning Model, which is a sub-model of the Hybrid Learning Approach, is especially in language classes. It has been concluded that this model, which has been applied in many countries, is extremely useful in language teaching and can be used efficiently especially in language preparation programs. It has been suggested to be applied at all levels of education in order to increase the academic success of the students and to encourage their participation in the lessons.

Keywords: Hybrid Teaching, Enriched Virtual Model, Distance Education, Educational Management.

1. INTRODUCTION

Today, due to the rapid development of science and technology and global health problems such as the Coronavirus disease 2019 (Corona Virus Disease, COVID-19), the use of technology has ceased to be a choice in many areas and has become a necessity. In particular, the COVID-19 pandemic has led to a rapid transition to the use of remote technologies in working and educational life. One of the areas where technology is most widely used and affected in this process has been in education such as language training, health science education and vocational education.

The knowledge, skills, attitudes and behaviors gained through education become more effective, efficient, systematic and holistic with the use of technology. Some of the traditional (in-class) methods and techniques commonly used in education include narration, question-answer, discussion, problem solving, observation, case analysis, role-playing, brainstorming, group work and educational games.

In cases where traditional education methods are insufficient, new methods and models are needed in education and training in parallel with technological developments. Distance education also emerges as an education system that responds to this need. In this study, it is aimed to reveal the concepts related to distance education, the examples of the use of distance education in language teaching, the methods and tools used in distance education.

Many educators in the educational innovation field are considering redesigning secondary and higher education to better reflect students' 21st century learning needs. Initiatives around the world such as the XQ Super School project and Ministries of Education's Next Generation High Schools are about restructuring the high school experience to include "real world" learning components as well as skills and passion-building opportunities such as internships,

* Corresponding author.

project-based learning, robust computer science curricula and competency-based learning systems. Of course, such experiences that foster deeper learning do not fit well with the current structure of traditional schools. Indeed, to integrate these unconventional practices into schools at large scale, massive logistical shifts in where, when, how and in what way a student learns become crucial. Traditional school structures and features, such as 45-minute lessons, age-based groups, and single-teacher classes, are becoming obsolete. Students no longer go to school every day, as new high school designs promote flexible, student-centred learning and even off-campus experiences. Looking to the future, educational theorists predict that not only will blended learning opportunities continue to increase in schools, but there will also be an inevitable rise of the Enriched Virtual Blended learning model at the High School level.

2. HYBRID LEARNING (BLENDED LEARNING)

The concept of blended learning has different definitions. Powell et al. (2015) argue that the blended learning approach brings together the best elements of online learning and face-to-face and that it is expected to become the most preferred model in the future and become much more common than other models.

In general terms, blended learning describes a learning experience that can be tailored to a particular audience and is not limited by place or time. With this in mind, we can express a broader definition as follows. It is the learning of a part of traditional education and the time allocated for it by controlling their own learning syllabus and speed at any time and place with the opportunities of the online environment. Blended learning as a concept is rooted in online learning and for a student traditional teaching never has; It represents a fundamental shift that has the potential to improve, develop and optimize student learning. (Maxwell 2016)

Although schools have long used computers and technology in teaching, until recently they have never used technology to provide students with a true blend of education that gives them control over their learning. In face-to-face education, technology is used in the classroom, but its use outside the classroom is more limited. In order for the students to understand the lesson better, activities and opportunities should be provided that will enable them to improve themselves at any time after the lesson. In addition, there was a need for applications that enable students to communicate with their teachers and friends about the lessons, to talk simultaneously and to exchange ideas. These opportunities are offered to learners by distance education. For this reason, distance learning, which offers more flexible learning environments and methods in terms of time and place, and the opportunities it provides to learning environments above, has been combined with face-to-face education.

Blended learning is primarily expressed as the mixing of web-based technologies (virtual classrooms, self-teaching education, collaborative learning, video, audio and text streaming) to achieve an instructional goal. Secondly, using a different conceptual approach, combining various educational approaches such as behaviorism, cognition, constructivism to produce the most appropriate learning outcomes with or without using instructional technologies is expressed as blended learning. Thirdly, using face-to-face education under the guidance of a teacher and combining different formats of instructional technologies (video, CD-Rom, web-based learning, distance learning technologies, etc.) is blended learning. Fourth and lastly, blended learning is defined as combining and blending real work tasks and learning technologies to create a harmonious effect for learning (Kayalar, 2021; Kayalar and Kayalar, 2020; Kayalar and Ağaoğlu 2020; Kazakoff et al, 2017; Schecter et al, 2015).

Blended learning does not have a limited scope, such as the use of some strategies such as discussion forums, mail, content presentation, which are used only in e-learning, in face-to-face teaching and mostly as a tool to support face-to-face teaching (Usta, 2007; Ünsal, 2010). Blended learning, which should be accepted as an instructional design approach, is a process that should be strategically planned in order to be implemented in the realization of teaching for a course, in the dimension of a curriculum or an educational institution.

Blended learning focuses on achieving the highest achievement by matching the right learning technologies and applying the learning objectives, with the right personal learning style to equip the right skills to the right person at the right time. The principles hidden in this definition include some important points such as

- The focus is on learning goals rather than sharing method.
- Many personal learning styles need support to reach large audiences.
- Each individual participates in the learning event with different information.
- In many cases, the most effective learning strategy is just what is needed at that moment.

There are main elements associated with blended learning such as psychological, technological, theoretical, communication and management system. Blended learning interacts with computer assisted learning, web based learning, electronic learning, learning management system and learning platform. In addition, there are other areas in which all these elements that blended learning is related to are also related to each other. For example, computer-assisted learning is directly related to behaviorism, computer-assisted applications, lifelong learning, and Web-assisted

learning, and developments here change by affecting computer-assisted learning. Web-supported or face-to-face learning approaches can be used in the school environment, inside and outside the classroom. One of them can be done or there may be a few of them. What is important here is the planning to be decided with the participation of other experts in the company of the instructor. In addition, it is the rational evaluation of the school's opportunities and strengths. The blended learning approach to be applied will have many benefits for both the student and the teaching staff, as well as the educational institution and its effectiveness. We can briefly explain some of the benefits of the blended learning approach as: it increases learning effectiveness and makes the richness of learning permanent, it is convenient in terms of time and cost, the esults take place at the most appropriate level, and collated studies occur immediately. In addition to these, the application of the blended learning approach includes learning wealth, access to information, social interaction, management of learning and so on. It leads to important results (Osguthorpe and Graham, 2003).

It is understood from these explanations that, in general, this learning approach provides diversity, cooperation and communication density in terms of individual learning, individual speed, listening, reading, seeing and application. In addition, it can be said that feedback has a positive effect such as speed, freedom in the learning environment, saving in time and learning costs. Of the sub-models of blended learning, A La Carte Learning Model and Enriched Virtual Learning Model have outstanding benefits and advantages in education during the recent Pandemic.

3. ENRICHED VIRTUAL LEARNING MODEL

Enriched Virtual model—a whole-school experience in which within each course (e.g., math), students divide their time between attending a brick-and-mortar campus and learning remotely using online delivery of content and instruction. Many Enriched Virtual programs began as full-time online schools and then developed blended programs to provide students with brick-and-mortar school experiences. The Enriched Virtual model differs from the Flipped Classroom because in Enriched Virtual programs, students seldom attend the brick-and-mortar campus every weekday. It differs from the A La Carte model because it is a whole-school experience, not a course-by-course model (Hunsinger, 2019).

According to Jenny White (2019) the Enriched Virtual model, by definition, is rather straightforward: the backbone of student learning is online and the student is only required to attend brick-and-mortar school on designated days. Thanks to technology, learning is happening anywhere, anytime, and (sometimes) at varying pace. The required face-to-face time within this model typically serves two main purposes:

1) enrich students' learning experiences with group-based work or teacher-led instruction, for example, and
2) hold students accountable via regular in-person check-ins with their teachers and advisor(s).

This model isn't flashy; it's quite operational in nature. In the Christensen Institute's research, early adoption of this model emerged among fully virtual schools that shifted to blended learning to provide stronger supports for students who otherwise struggle to stay on track. It's worth noting, however, that this blended model is considered disruptive: it provides learning opportunities not governed by seat time like traditional instructional models, but instead determined by the degree to which students control time, path, pace, and in some cases, place of their learning.

Today, this model is still fairly niche among traditionally brick-and-mortar public schools; few schools tout their "Enriched Virtual" model at conferences or in articles. But when you take a look at schools leveraging this model, their Enriched Virtual practice is pretty exciting in that it has the potential to benefit an increasing number of today's students: without Enriched Virtual's disruptive blended-learning structure, scheduling flexibility, off-campus learning experiences, opportunities to explore emerging passions, and more, wouldn't be feasible for schools to provide.

Many reasons for choosing an enriched virtual model or an à la carte model are the same. Both models allow you to support student-centred learning, develop self-management skills, and personalize learning through a wider range of course options. They can be used to accelerate credit accumulation, resolve timing constraints, or support basic learning skills. Often these models are used with "non-traditional" students. For example, over-age under-credited high school students that have had interrupted academic progress, may need courses that don't fit into their traditional schedule. These models also help in rural areas where some students commute to school too long.

The course needs should be identified at the school in order to get started with the a la carte model or enriched model. Replies for these questions should be found;

- Are there gaps in your course offerings?
- Have students become disengaged in the required courses?
- What are the interests?
- Do you have students who are far below grade level and need an additional course to meet their needs?
- Who will monitor student progress?

Both the a la carte and enriched virtual models are closer to online learning in the spectrum of blended learning and are more often used in higher grade levels. They are classified as blended learning because they still include limited face-to-face time with a teacher. An "a la carte model" often has mostly traditional face to face classes with an online course supplement whereas the enriched virtual model is mostly online with intermittent face-to-face interactions.

Blended Learning

- Students learn in part through online learning and are in control of where, when and how they study during their learning.
- Devices are used to take advantage of the individualization opportunity.
- Classrooms are fundamentally changing teaching to provide an integrated learning experience.

Technology Assisted Learning

- Students use technology to do the same job at the same place, time and speed while learning.
- Devices are used to support traditional teaching
- Classes are supplemented to enrich the traditional learning experience.

CONCLUSION

Distance education, which has become widespread in our country and in the world with the COVID-19 pandemic, can be offered synchronously (interactively) or asynchronously (non-interactive), without time and place restrictions. It is seen that distance education is not sufficient in subjects such as clinical applications and laboratory modules for applied sciences such as Language Education and Science Education. For this reason, a mixed education model in which online and face-to-face methods are used together is recommended for applied sciences (Kozan et al, 2021). When the national and international literature studies in which the blended education model is used in Language Education are examined, it is seen that Enriched Virtual Learning Model, which is a type of blended education, supports pedagogy and positively affects the learning outcomes of the students, so it is more recommended than other models. It is recommended to use and disseminate tools such as online course materials, discussion boards, simulation, videos, mobile technologies and social media platforms to facilitate learning in virtual language teaching.

References

[1] Kayalar, F. & Ağaoğlu, A. (2020). The Importance of Blended Learning During the Period Of Global PandemicCurrent Researches in Educational Sciences, Editor: Assist Prof. Mustafa Batuhan Kurt. Duvar Publishing, İzmir

[2] Kayalar, F.& Kayalar, F. (2020). Karma Eğitim Modelinde Ters-yüz Öğrenme. Eğitim Bilimleri Alanında Güncel Araştırmalar, Editör: Dr. Öğr. Üyesi Mustafa Batuhan Kurt, Duvar Yayınları, İzmir

[3] Kayalar, M. T. (2021). The Efficiency of Flipped Classroom Model in Virtual Education (Chapter 4). Academic Research and Reviews in Educational Sciences Editor: Assoc. Prof. Aydın Kızılaslan. Duvar Publishing, İzmir.

[4] Hunsinger, J. (2019) Blended Learning, Curriculum & Instruction, Instructional Technology, K-12 Teachers, Professional Development for Teachers, Resources, Teacher Tips, Jul 30, 2019

[5] Kazakoff, E. R., Macaruso,P.& Hook, P. (2017). Efficacy of a blended learning approach to elementary school reading instruction for students who are English Learners. Education Tech Research Dev. https://doi.org/10.1007/s11423-017-9565-7

[6] Kozan, E.H., Çolak, M. & Demirhan, B.S. (2021). Distance Education in COVID-19 Pandemia: Reflections on Nursing Education.. J Educ Res Nurs. 2021;18(Supp. 1).

[7] Osguthorpe, R. T. ve Graham, C. R. (2003). Blended learning environments: Definitions and directions. Quarterly Review of Distance Education, 4, 227-233.

[8] Powell, A., Watson, J., Staley, P., Patrick, S., Horn, M., Fetzer, L., Hibbard, L., Oglesby, J., & Verma, S. (2015). Blending Learning: The evolution of online and face-to-face education from 2008–2015. Vienna, VA: International Association for K–12 Online Learning. Retrieved from http://www.inacol.org/wp-content/uploads/2015/07/iNACOL_Blended-Learning-The-Evolution-of-Online-And-Face-toFace-Education-from-2008-2015.pdf.

[9] Schechter, R., Macaruso, P., Kazakoff, E. R., & Brooke, E. (2015). Exploration of a blended learning approach to reading instruction for low SES students in early elementary grades. Computers in the Schools, 32, 183–200

[10] Thorne, K. (2003). Blended learning: how to integrate online and traditional learning, London: Kogan Page

[11] Usta, E. (2007). Harmanlanmış Öğrenme ve Çevrimiçi Öğrenme Ortamlarının Akademik Başarı ve Doyuma Etkisi. Yayınlanmamış Doktora Tezi, Gazi Üniversitesi, Ankara.

[12] Ünsal, H. (2010). Yeni bir öğrenme yaklaşımı: Harmanlanmış öğrenme. Milli Eğitim, 185, 130-137.

[13] White, J. (2019) Jenny White Blended learning Model https://www.blendedlearning.org/is-the-enriched-virtual-blended-learning-model-the-future-of-high-school/)

[14] Maxwell, C. (2016). What blended learning is-and isn't. BLU: Blended Learning Universe. Erişim adresi: http://www.blendedlearning.org/what-blended-learning-is-and-isnt/

Engineering, Transport, IT and Artificial Intelligence (IAC-ETITAI 2021)

Identification of Deceptive Opinion Spam Based on Deep Neural Networks

Aytuğ ONAN[a]1

[a] *İzmir Katip Çelebi University, Faculty of Engineering and Architecture, Department of Computer Engineering, İzmir/Turkey, aytug.onan@ikcu.edu.tr*

Abstract

The number of review texts shared on the web for products and services is increasing significantly. Review texts inform readers about the positives and negative aspects of a particular product, and both individual and institutional decision makers rely on them. Deceptive content is created for two primary reasons. The first method is to create a positive fictitious review text for a specific product or service. The second category is intended to purposefully diminish the dignity of a specific product or service. Detecting and filtering undesirable content is a critical application area for the Web. In this paper, we present a deep learning-based approach for identification of deceptive opinion spam. In the empirical analysis, four word embedding schemes (i.e., word2vec, GloVe, fastText, and LDA2vec) have been considered. The predictive performances of neural language models have been compared in conjunction with convolutional neural network, recurrent neural network, long short-term memory, and gated recurrent units. The experimental results indicate that gated recurrent units yield promising results for deceptive opinion spam identification.

Keywords: opinion spam, deep learning, machine learning, review texts

1. INTRODUCTION

The Web has become an integral part of electronic commerce and marketing because of advancements in information technology. The number of review texts shared on the web for products and services is growing rapidly. Review texts inform readers about the positive and negative aspects of a product, and both individual and institutional decision makers rely on them.

Review texts serve as a vital information resource for both commercial and non-commercial organizations. Individual decision-makers and business organizations rely on review texts to identify and eliminate product quality weaknesses [1]. Positive attitudes toward a particular product, service, or individual result in the relevant product, service, or individual gaining a competitive advantage. As a result, the power of user-generated content on social media may result in the creation of content that bears the characteristics of an advertisement for a specific product or service, lacks information about the product or service, and portrays the product's characteristics in a falsely negative or positive light [2, 3].

Deceptive content is generated for primarily two reasons. The first is to fabricate a favourable false review text for a specific product or service. The second category is designed to purposefully diminish the dignity of a specific product or service. Deceptive opinion spam identification and filtering is a critical application area for the Web. The conventional supervised learning models, including Naïve Bayes, k-nearest neighbour, and support vector machines have been successfully employed in areas, such as spam filtering [3, 4]. Although it encompasses sub-problems such as identifying unwanted review texts, sharing unrealistic opinions, sharing content exclusively for a specific brand, or

* Corresponding author.

sharing content that does not qualify as a review, determining unrealistic opinions appears to be the most challenging [5].

The purpose of this study is to evaluate deep neural networks' predictive performance in identifying deceptive opinion spam. In this regard, we considered schemes for word embedding (namely, word2vec, GloVe, fastText, and LDA2vec). The predictive abilities of neural language models were evaluated in comparison to those of convolutional neural networks, recurrent neural networks, long short-term memory, and gated recurrent units. The experimental results indicate that gated recurrent units can be used to identify deceptive opinion spam.

This paper is divided into five sections. Section 2 discusses related work on deceptive opinion spam classification. Section 3 discusses the study's methodology, including the dataset, word embedding schemes, and deep learning architectures. Section 4 details the experimental design and methodology, as well as the empirical findings. Finally, in Section 5, the concluding remarks are presented.

2. RELATED WORK

Numerous studies have been conducted on determining review spam texts using machine learning and deep learning methods. This section summarizes some of the prior works in this field.

Jindal and Liu [2] discussed two classification problems, namely identifying duplicate review texts, and determining undesirable review texts. Classification algorithms such as Naïve Bayes and logistic regression are used. In another study, Jindal and Liu [6] classified the problem of identifying unfavorable review texts into three categories: unrealistic evaluations, brand-specific evaluations, and texts that do not qualify for review. In this study, supervised learning algorithms based on machine learning were used to identify deceptive opinion spam across all three categories. Lim et al. [7] proposed a method for modelling voting behaviors to identify unfavorable review texts. The purpose of this study is to simulate the behavior of individuals who create unwanted content in the direction of supporting a particular product or service. Algur et al. [8] developed a conceptual method for identifying review spam based on similarity measures. Ott et al. [9] conducted a comprehensive evaluation of the effectiveness of various feature sets for the problem of identifying review spam texts. The results of this study are presented in comparison to human achievements because of training these attributes using supervised learning methods and a variety of text attributes, such as type determination-based attributes, linguistic attributes, 1-gram, 2-gram, and 3-gram. The study found that when 2-gram and linguistic features were combined, the best performance was obtained. In a study conducted by [10], undesirable evaluation texts were identified using frequent pattern mining techniques. Another study [11] used text analysis to identify review spam texts that contained only negative opinions. Narayan et al. [12] has developed a machine learning-based method for identifying review spam texts. In this study, decision trees, Naïve Bayes, support vector machines, k-nearest neighbor, and logistic regression classifiers were used to evaluate linguistic features. In another study, Onan [3] examined the predictive performance of psychological and linguistic feature sets and their ensemble combinations for identification of deceptive opinion spam. The predictive performance of feature sets has been examined by supervised learning algorithms, including Naïve Bayes algorithm, k-nearest neighbor algorithm, support vector machines, and C4.5 algorithm.

Recently, deep neural networks have been also employed for review spam identification. For instance, Jain et al. [13] presented a hybrid architecture which combines convolutional neural networks and long short-term memory architecture for opinion spam identification. The presented scheme achieved promising predictive performance on the task, with a classification accuracy of 95%. Similarly, Baccouche et al. [14] presented a deep learning approach to identify malicious texts from public comments and emails. In another study, Fahfouh et al. [15] introduced a novel deep neural network architecture to identify deceptive opinion spam. In this scheme, a denoising autoencoder and a paragraph vector model have been concatenated.

In another study, Zhang et al. [16] presented a recurrent convolutional neural network architecture for deceptive review identification. In this scheme, each word in a review text has been represented as a recurrent convolutional vector to distinguish between deceptive and true contextual knowledge embodied in online reviews. Bhuvaneshwari et al. [17] also presented a hybrid deep neural network architecture based on convolutional neural networks and bidirectional long short-term memory architecture for deceptive opinion spam identification. In another study, Wang et al. [18] presented an attention-based neural network that identifies spam reviews by utilizing linguistic and behavioral feature sets. In this scheme, multilayer perceptron is employed to extract behavioral features, whereas convolutional neural network has been utilized to extract linguistic features. In addition, the attention mechanism has been built on top of that layer and has been designed to identify deceptive opinion spam. In another study, Anass et al. [19] examined the predictive performance of deep neural networks for deceptive opinion spam identification. In this study, convolutional neural network, recurrent neural network, long short-term memory, bidirectional long short-

term memory, gated recurrent unit, and bidirectional gated recurrent unit has been evaluated in terms of accuracy, F-measure, precision, and recall.

Text classification is a critical subfield of text mining that involves classifying a text document into one or more predefined categories or classes [20]. Text mining has been applied successfully in a variety of fields, including web page classification [21], sentiment analysis [22-27], and text genre classification [28]. It is critical to filter out deceptive opinion spam using an appropriate text representation and deep learning architecture. This study evaluates the efficacy of conventional word embedding methods and deep learning architectures in this context.

3. METHODOLOGY

This section introduces deceptive opinion spam data set, the word embedding methods used to represent text documents, and deep learning architectures employed in the empirical analysis.

3.1. Dataset

To examine the predictive performance of word embedding methods and deep learning architectures on deceptive opinion spam, we have utilized the dataset referred as "Deceptive Opinion Spam Corpus" constructed by Ott et al. [9, 11]. This dataset contains legitimate and fabricated review texts for twenty hotels in Chicago. There are 1600 evaluation texts in the collection. 800 of these texts are genuine evaluations, while 800 are deemed unsuitable for evaluation. 400 of the actual evaluation texts express favorable sentiments, while 400 express negative sentiments. Similarly, 400 of the undesirable review texts express positive sentiments, while 400 express negative sentiments [9, 11].

3.2. Neural language models

The bag-of-words (BOW) method, which is frequently used to represent text documents in natural language processing tasks, has two significant drawbacks. The order of the words is irrelevant in BOW. Additionally, the vector length is equal to the total number of words in the text collection. Word embedding methods, which allow for the representation of words in text documents in dense spaces using fixed-length vectors, enable for the representation of data with fewer dimensions and are frequently used in deep learning-based text mining applications [29, 30]. In the empirical analysis, we have considered four schemes for word embedding (i.e., word2vec, GloVe, fastText, and LDA2vec).

The Word2vec model is a three-layer method for word embedding based on artificial neural networks [31]. It consists of an input layer, an output layer, and a hidden layer. Its objective is to learn word embedding by calculating the probability that a given word will be surrounded by other words. The model is composed of two fundamental architectures: skip-gram (SG) and continuous-bag-of-words (CBOW). The CBOW architecture defines the target word by analyzing the content of each word; on the other hand, the SG architecture takes the target word as input and predicts the words that surround it. The CBOW architecture can work effectively with small amounts of data. On the other hand, the SG architecture performs better with large data sets.

GloVe is a word2vec-based model that was developed to efficiently learn word embedding from text documents. The model combines the word2vec model's local contextual learning with global matrix factorization [32]. The probability ratios of words are also considered when calculating the error function in the model. In the learning process, words that are observed close together in a text document and are more likely to be seen together are more important than other words [31].

Another effective method for extracting word embeddings from text documents is the FastText model. Each word is represented in this model by dividing the character into n-grams. For each n-gram in the training set, word vectors are created. For morphologically rich languages and rare words, the fastText model provides a more efficient word embedding scheme [33].

The LDA2vec model is another method for word embedding that is based on the word2vec model and hidden Dirichlet separation. Dense word vectors are combined with hidden document level vectors in this method, depending on the Dirichlet distribution. The model enables the definition of topic-specific text collections and word vectors. In

this context, the model provides a topic-enriched word embedding scheme by associating each word with a topic. The model's objective function is skip-gram negative sampling [6].

3.3. Deep neural network architectures

The predictive performances of neural language models were evaluated in comparison to those of convolutional neural networks, recurrent neural networks, long short-term memory, and gated recurrent units. This section summarizes deep neural network architectures.

Convolutional neural networks (CNN) are a subset of deep neural networks that utilize a grid-based topology to process data [35]. Convolution, a specialized type of mathematical operation, was used in one or more convolutional layers of CNN, rather than general matrix multiplication, as is the case with conventional neural network architectures. The CNN architecture is composed of three layers: input, output, and hidden. The architecture's hidden layers serve in place of several layers, including convolutional layers, pooling layers, fully connected layers, and normalization layers. Convolution has been used on the input data in convolutional layers. The feature maps have been extracted as a result. To add nonlinearity to the architecture, activation functions (such as rectified linear unit) were used in conjunction with feature maps. Following convolution, pooling layers combine the outputs of neuron clusters, gradually shrinking the spatial size of the feature space. As a result, the model's parameter count has been reduced and its ability to avoid overfitting has been improved. Maximum pooling scheme is a common function in the pooling layer. The maximum value from each cluster was used in this scheme. Following convolutional and pooling layers, fully connected layers produce the architecture's final output.

Recurrent neural network (RNN) is a deep network architecture used to process sequential data. In this architecture, the connections between nerve cells form a directed diagram. The RNN architecture enables any length of input to be processed. During the calculation, historical information is also considered [36]. In RNN architectures, the exploding/vanishing gradient problem is often encountered. Modelling of long-term dependencies in RNN architectures is difficult due to the multiplicative gradient that increases or decreases based on the number of layers [37].

Long short-term memory units (LSTM) are a recursive neural network architecture to avoid the vanishing/exploding gradient problem encountered in traditional RNN architectures. LSTM allows the error to propagate backwards with a limited number of time steps. A basic LSTM unit consists of a cell and three basic gates: input gate, output gate and forget gate. According to the status of the gates, it is determined which information should be protected and when the units will be accessed [37].

The gated recurrent unit (GRU) is a recurrent neural network architecture that has a less complex architecture compared to LSTM and gives similar experimental results [36]. A basic GRU architecture has two gates, the reset gate, and the update gate.

4. EXPERIMENTAL PROCEDURE AND RESULTS

This section introduces the experimental procedure, the evaluation measures, and the empirical results.

4.1. Experimental procedure

TensorFlow and Keras were used to develop and train the deep learning architectures used in the empirical analysis. To optimize the predictive performance of each deep learning model, we used a hyper parameter searching algorithm. This was accomplished using hyper parameter optimization based on Bayesian optimization using the Gaussian process. For the corpus, 80% of data was used as the training set, while the remaining 20% was used as the testing set. The word2vec and fastText schemes, as well as the continuous skip-gram and CBOW schemes, were considered with varying vector sizes (200 and 300) and projection layer dimensions (100 and 200). Several parameters (including the number of topics and the negative sampling exponent) have been considered when developing the LDA2vec scheme. For LDA2vec scheme, number of topics has been taken as (N = 25) and the negative sampling exponent has been taken as ($\beta \in 0.75$).

4.2. Evaluation measures

To evaluate the performance of classification algorithms, two different evaluation measures, namely, classification accuracy and F-measure have been considered.

Classification accuracy (ACC) is the proportion of true positives and true negatives obtained by the classification algorithm over the total number of instances as given by Equation 1:

$$ACC = \frac{TN + TP}{TP + FP + FN + TN} \quad (1)$$

where *TN* denotes number of true negatives, *TP* denotes number of true positives, *FP* denotes number of false positives and *FN* denotes number of false negatives.

Precision (PRE) is the proportion of the true positives against the true positives and false positives as given by Equation 2:

$$PRE = \frac{TP}{TP + FP} \quad (2)$$

Recall (REC) is the proportion of the true positives against the true positives and false negatives as given by Equation 3:

$$REC = \frac{TP}{TP + FN} \quad (3)$$

F-measure takes values between 0 and 1. It is the harmonic mean of precision and recall as determined by Equation 4:

$$F - measure = \frac{2 * PRE * REC}{PRE + REC} \quad (4)$$

4.3. Experimental results and discussion

In Tables 1-2, the classification accuracy and F-measure values obtained by the compared models have been presented, respectively. In Figures 1-2, the main effects plots for classification accuracy and F-measure values have been presented to summarize the main findings of the empirical results.

Table 1. Classification accuracy values obtained by compared word embedding schemes and deep neural networks

Word embedding	Vector size	Layer size	CNN	RNN	LSTM	GRU
word2vec (Skip-gram)	200	100	69.38	76.16	77.91	80.31
word2vec (Skip-gram)	200	200	69.53	76.17	78.07	80.31
word2vec (Skip-gram)	300	100	70.82	76.34	78.18	80.38
word2vec (Skip-gram)	300	300	70.87	76.48	78.28	80.42
word2vec (CBOW)	200	100	73.25	76.92	78.86	81.02
word2vec (CBOW)	200	200	73.54	77.07	78.90	81.33
word2vec (CBOW)	300	100	73.56	77.29	79.07	81.40
word2vec (CBOW)	300	300	73.65	77.32	79.09	81.51
fastText (Skip-gram)	200	100	71.08	76.49	78.51	80.63
fastText (Skip-gram)	200	200	71.27	76.74	78.53	80.72
fastText (Skip-gram)	300	100	72.60	76.85	78.59	80.80
fastText (Skip-gram)	300	300	73.22	76.92	78.66	80.86
fastText (CBOW)	200	100	75.03	77.48	79.45	83.29
fastText (CBOW)	200	200	75.36	77.52	79.47	83.65
fastText (CBOW)	300	100	75.49	77.56	79.71	84.74
fastText (CBOW)	300	300	75.59	77.62	79.76	84.85
GloVe	200	100	75.73	77.63	79.77	85.15
GloVe	200	200	75.88	77.65	79.99	85.49
GloVe	300	100	76.02	77.69	80.13	85.76
GloVe	300	300	76.13	77.79	80.28	86.08
LDA2vec	200	100	74.10	77.38	79.18	81.58
LDA2vec	200	200	74.20	77.40	79.19	82.53
LDA2vec	300	100	74.86	77.44	79.26	82.56
LDA2vec	300	300	75.01	77.48	79.41	82.60

Regarding the classification accuracy values obtained by the neural language models and deep neural network architectures, we have considered six word embedding configurations and four deep learning architectures. As can be seen from the empirical results listed in Table 1, Glove word embedding scheme outperforms the other neural language

models. The second highest predictive performances have been achieved by fastText (CBOW) model. The third highest predictive performances have been achieved by LDA2vec scheme, which is followed by word2vec (CBOW) model. As can be seen from the empirical results, CBOW models considered in conjunction with fastText and word2vec outperform skip-gram models considered in conjunction with fastText and word2vec. Regarding the vector sizes considered in the experimental analysis, the vector size of 300 outperforms the vector size of 200. In addition, three different values have been considered for dimension of projection layer. Regarding the dimension of projection layer, layer size of 300 outperforms the other compared configurations. Regarding the performance of deep neural network architectures in the empirical analysis, gated recurrent unit (GRU) outperforms all the other models. The second highest predictive performances have been achieved by long short-term memory (LSTM) architecture, which is followed by recurrent neural network (RNN) architecture. The highest predictive performance for deceptive opinion spam identification has been achieved by GloVe word embedding scheme in conjunction with GRU, with a classification accuracy of 86.08%. The experimental patterns that are valid for classification accuracy hold true for F-measure values as well.

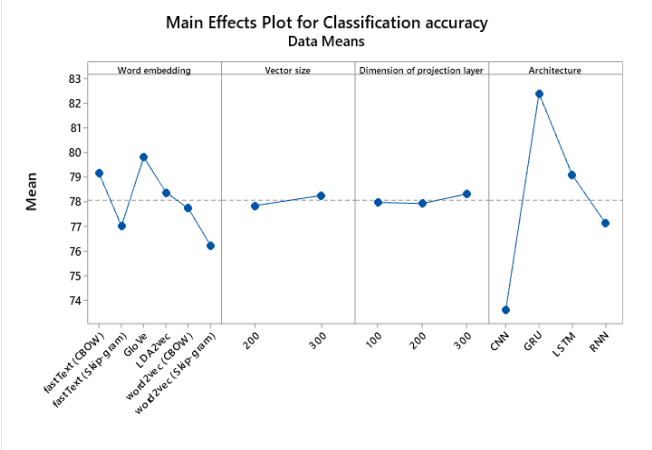

Fig. 1. Main effects plot for classification accuracy values

Table 2. Classification accuracy values obtained by compared word embedding schemes and deep neural networks

Word embedding	**Vector size**	**Layer size**	**CNN**	**RNN**	**LSTM**	**GRU**
word2vec (Skip-gram)	200	100	0.70	0.77	0.79	0.81
word2vec (Skip-gram)	200	200	0.70	0.77	0.79	0.81
word2vec (Skip-gram)	300	100	0.71	0.77	0.79	0.81
word2vec (Skip-gram)	300	300	0.71	0.77	0.79	0.81
word2vec (CBOW)	200	100	0.74	0.78	0.80	0.82
word2vec (CBOW)	200	200	0.74	0.78	0.80	0.82
word2vec (CBOW)	300	100	0.74	0.78	0.80	0.82
word2vec (CBOW)	300	300	0.74	0.78	0.80	0.82
fastText (Skip-gram)	200	100	0.72	0.77	0.79	0.81
fastText (Skip-gram)	200	200	0.72	0.77	0.79	0.81
fastText (Skip-gram)	300	100	0.73	0.78	0.79	0.81
fastText (Skip-gram)	300	300	0.74	0.78	0.79	0.82
fastText (CBOW)	200	100	0.76	0.78	0.80	0.84
fastText (CBOW)	200	200	0.76	0.78	0.80	0.84
fastText (CBOW)	300	100	0.76	0.78	0.80	0.85
fastText (CBOW)	300	300	0.76	0.78	0.80	0.86
GloVe	200	100	0.76	0.78	0.80	0.86
GloVe	200	200	0.77	0.78	0.81	0.86
GloVe	300	100	0.77	0.78	0.81	0.86
GloVe	300	300	0.77	0.78	0.81	0.87
LDA2vec	200	100	0.75	0.78	0.80	0.82
LDA2vec	200	200	0.75	0.78	0.80	0.83
LDA2vec	300	100	0.76	0.78	0.80	0.83
LDA2vec	300	300	0.76	0.78	0.80	0.83

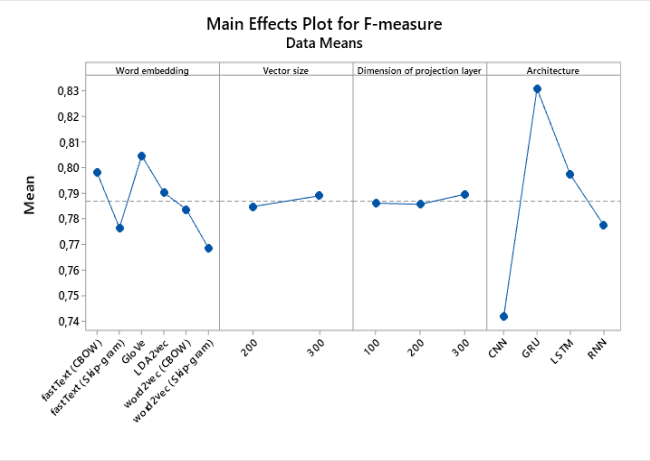

Fig. 2. Main effects plot for F-measure values

CONCLUSION

The number of product and service review texts shared on the web is increasing significantly. Review texts inform readers about the advantages and disadvantages of a particular product, and they are relied upon by both individual and institutional decision makers. Two primary reasons are given for the creation of deceptive content. The first technique is to create a fictitious positive review text for a particular product or service. The second category is intended to purposefully diminish a product's or service's dignity. Unwanted content detection and filtering is a critical application area for the Web. We present a deep learning-based approach for detecting deceptive opinion spam in this article. Four word embedding schemes were considered in the empirical analysis (word2vec, GloVe, fastText, and LDA2vec). The predictive abilities of neural language models were evaluated in comparison to those of convolutional neural networks, recurrent neural networks, long short-term memory, and gated recurrent units. The highest predictive performance for deceptive opinion spam identification has been achieved by GloVe word embedding scheme in conjunction with GRU, with a classification accuracy of 86.08%. The experimental results indicate that gated recurrent units can be used to identify deceptive opinion spam.

References

[1] A Y. Lin, T. Zhu, X. Wang, J. Zhang, and A. Zhou, "Toward online review spam detection"," in Proc. of the 23rd International Conference on World Wide Web, 2014, pp. 341–342.
[2] N. Jindal and B. Liu, "Review spam detection," in Proceedings of the 16th international conference on World Wide Web - WWW '07, 2007.
[3] A. Onan, "Review spam detection based on psychological and linguistic features," in 2018 26th Signal Processing and Communications Applications Conference (SIU), 2018.
[4] H. Drucker, D. Wu, and V. N. Vapnik, "Support vector machines for spam categorization," IEEE Trans. Neural Netw., vol. 10, no. 5, pp. 1048–1054, 1999.
[5] B. Liu, Sentiment Analysis and Opinion Mining. San Rafael, CA: Morgan & Claypool, 2012.
[6] N. Jindal and B. Liu, "Opinion spam and analysis," in Proceedings of the international conference on Web search and web data mining - WSDM '08, 2008.
[7] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, "Detecting product review spammers using rating behaviors," in Proceedings of the 19th ACM international conference on Information and knowledge management - CIKM '10, 2010.
[8] S. P. Algur, A. P. Patil, P. S. Hiremath, and S. Shivashankar, "Conceptual level similarity measure based review spam detection," in 2010 International Conference on Signal and Image Processing, 2010.
[9] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," arXiv [cs.CL], 2011.
[10] A. Mukherjee, B. Liu, and N. Glance, "Spotting fake reviewer groups in consumer reviews," in Proceedings of the 21st international conference on World Wide Web - WWW '12, 2012.
[11] M. Ott, C. Caride, and J. T. Hancock, "Negative Deceptive Opinion Spam," in Proceedings of NAACL-HLT 2013, pp. 497–501.
[12] R. Narayan, J. K. Rout, and S. K. Jena, "Review spam detection using opinion mining," in Advances in Intelligent Systems and Computing, Singapore: Springer Singapore, 2018, pp. 273–279.
[13] G. Jain, M. Sharma, and B. Agarwal, "Spam detection in social media using convolutional and long short term memory neural network," Ann. Math. Artif. Intell., vol. 85, no. 1, pp. 21–44, 2019.
[14] A. Baccouche, S. Ahmed, D. Sierra-Sosa, and A. Elmaghraby, "Malicious text identification: Deep Learning from public comments and emails," Information (Basel), vol. 11, no. 6, p. 312, 2020.

[15] A. Fahfouh, J. Riffi, M. Adnane Mahraz, A. Yahyaouy, and H. Tairi, "PV-DAE: A hybrid model for deceptive opinion spam based on neural network architectures," Expert Syst. Appl., vol. 157, no. 113517, p. 113517, 2020.
[16] W. Zhang, Y. Du, T. Yoshida, and Q. Wang, "DRI-RCNN: An approach to deceptive review identification using recurrent convolutional neural network," Inf. Process. Manag., vol. 54, no. 4, pp. 576–592, 2018.
[17] P. Bhuvaneshwari, A. N. Rao, and Y. H. Robinson, "Spam review detection using self attention based CNN and bi-directional LSTM," Multimed. Tools Appl., 2021.
[18] X. Wang, K. Liu, and J. Zhao, "Detecting deceptive review spam via attention-based neural networks," in Natural Language Processing and Chinese Computing, Cham: Springer International Publishing, 2018, pp. 866–876.
[19] F. Anass, R. Jamal, M. A. Mahraz, Y. Ali, and H. Tairi, "Deceptive opinion spam based on deep learning," in 2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS), 2020.
[20] A. Onan, S. Korukoğlu, and H. Bulut, "Ensemble of keyword extraction methods and classifiers in text classification," Expert Syst. Appl., vol. 57, pp. 232–247, 2016.
[21] A. Onan, "Classifier and feature set ensembles for web page classification," J. Inf. Sci., vol. 42, no. 2, pp. 150–165, 2016.
[22] A. Onan, S. Korukoğlu, and H. Bulut, "A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification," Expert Syst. Appl., vol. 62, pp. 1–16, 2016.
[23] A. Onan and S. Korukoğlu, "A feature selection model based on genetic rank aggregation for text sentiment classification," J. Inf. Sci., vol. 43, no. 1, pp. 25–38, 2017.
[24] A. Onan, "Hybrid supervised clustering based ensemble scheme for text classification," Kybernetes, vol. 46, no. 2, pp. 330–348, 2017.
[25] A. Onan and M. A. Tocoglu, "A term weighted neural language model and stacked bidirectional LSTM based framework for sarcasm identification," IEEE Access, vol. 9, pp. 7701–7722, 2021.
[26] A. Onan, S. Korukoğlu, and H. Bulut, "A hybrid ensemble pruning approach based on consensus clustering and multi-objective evolutionary algorithm for sentiment classification," Inf. Process. Manag., vol. 53, no. 4, pp. 814–833, 2017.
[27] M. A. Toçoğlu and A. Onan, "Satire detection in Turkish news articles: A machine learning approach," in Communications in Computer and Information Science, Cham: Springer International Publishing, 2019, pp. 107–117.
[28] A. Onan, "An ensemble scheme based on language function analysis and feature engineering for text genre classification," J. Inf. Sci., vol. 44, no. 1, pp. 28–47, 2018.
[29] A. Onan, "Mining opinions from instructor evaluation reviews: A deep learning approach," Comput. Appl. Eng. Educ., vol. 28, no. 1, pp. 117–138, 2020.
[30] A. Onan, "Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks," Concurr. Comput., 2020.
[31] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," arXiv [cs.CL], 2013.
[32] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014.
[33] S. Sukheja, S. Chopra, and M. Vijayalakshmi, "Sentiment Analysis using Deep Learning – A survey," in 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), 2020.
[34] C. E. Moody, "Mixing Dirichlet topic models and word embeddings to make lda2vec," arXiv [cs.CL], 2016.
[35] D. Ciresan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2012.
[36] P. Baby and Krishnapriya, "Sentimental Analysis and deep learning : A survey," Int. J. Sci. Res. Sci. Eng. Technol., pp. 212–220, 2020.
[37] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural Language Processing (almost) from Scratch," arXiv [cs.LG], 2011.

Deep Learning Based Identification of Offensive and Hateful Language on Social Media

Aytuğ Onan[a]1

[a] *İzmir Katip Çelebi University, Faculty of Engineering and Architecture, Department of Computer Engineering, İzmir/Turkey, aytug.onan@ikcu.edu.tr*

Abstract

While the Internet and mobile devices have greatly simplified the access to social media platforms, this in turn facilitates the transfer of ideas, opinions, and feelings, but such platforms are subject to misuse and offensive and hateful language can also be shared easily on social media, as well. The unregulated expansion of hate has the potential to severely damage our society, and it also has the potential to damage marginalized groups and individuals. In social media, the practice of reviewing content, both manually and algorithmically, to filter out problematic language and content has become prevalent. The manual scheme is bound to become increasingly inefficient as the number of users and their posts for social media platforms both progressively increase. Hence, the identification of offensive and hateful language on social media is a crucial task in natural language processing. In this paper, we evaluate the predictive performance of deep neural networks for identification of offensive and hateful language. Three different word embedding schemes have been studied in the empirical analysis: word2vec, GloVe, and fastText. Convolutional neural networks, recurrent neural networks, long short-term memory, gated recurrent unit, bidirectional long short-term memory and bidirectional gated recurrent units have all been examined. The empirical results indicate that bidirectional gated recurrent units yield promising results for the task.

Keywords: offensive text, hateful language, deep learning, bidirectional gated recurrent unit.

1. INTRODUCTION

Advances in information and communication technologies, as well as the proliferation and ubiquity of mobile devices equipped with Internet access, have significantly increased access to social media platforms. As a result, the number of social media users and the content they share on social media platforms, such as feelings, thoughts, and opinions, continues to grow. However, easy access to social media platforms can also result in abuse, as it enables the rapid dissemination of hateful or harassing content.

Online social media platforms make it simple and global to express oneself, whether it is through normal online discourse or hate speech. Typically, hate speech is directed at an individual, a group, or an organization, consists of insulting statements intended to incite hatred or harm one's reputation [1]. A critical component of today's online social platforms is the automatic detection of hate and offensive expressions in natural languages, which are quite peculiar in their nature and frequently difficult to detect [2]. There are several important reasons why it is difficult to detect hateful and/or offensive speech automatically, particularly on social media. There are many expressions that are not offensive in nature, but they may contain abusive or hateful connotations in the context in which they are used [3]. In addition, words can be deliberately used differently to circumvent automatic content checking [4].

* Corresponding author.

Certain platforms, including widely used social media platforms such as Twitter and Facebook, employ human labor to attempt to remove malicious, abusive, offensive, or hateful content [1]. However, due to the volume of content shared on these platforms, it is extremely difficult to manually remove irrelevant content. This highlights the critical nature of automatically detecting hate speech, harassment, malicious, or offensive content in content shared on social media platforms.

The main goal of this study is to investigate how effective deep neural networks are in identifying offensive and hateful language on social media. For this reason, we have examined several approaches for embedding words (namely, word2vec, GloVe, and fastText). Convolutional neural networks, recurrent neural networks, long short-term memory, gated recurrent unit, bidirectional long short-term memory and bidirectional gated recurrent units have all been examined. The empirical results indicate that bidirectional gated recurrent units yield promising results for the task.

This paper is made up of five parts. The second section of this paper discusses previous work on offensive and hateful speech identification. Section 3 details the methodology for the study, including the dataset, word embedding schemes, and deep learning architectures. Section 4 includes the details of the experiment, along with the empirical results. Finally, Section 5 presents the conclusion.

2. RELATED WORK

Much attention has been paid to hate speech detection because of the variety of tasks it can be applied to. Much of the earlier research contributions have employed linguistic features alongside supervised machine learning techniques.

Davidson et al. [5] examined the predictive performance of conventional text representation schemes and supervised learning methods for automated identification of hate speech. The empirical results indicated that TF-IDF weighted n-gram models in conjunction with logistic regression algorithm can yield promising results for hate speech recognition. Sharma et al. [6] developed an ontological classification scheme for identifying harmful speech based on the degree of hateful intent and used it to annotate Twitter data. In another study, Waseem and Hovy [7] examined the predictive performance of demographic, lexical, and geographic features on hate speech recognition. The empirical results indicated that geographic features and features based on word length distributions does not improve the predictive performance of supervised learning algorithms. However, gender information in conjunction with character n-grams can be utilized as an effective feature set for hate speech recognition task. Similarly, Nobata et al. [4] employed n-gram features, linguistic features, syntactic features, and distributional semantics features for abusive language identification in online user content.

Agarwal and Sureka [8] conducted another study in which they used topic modeling, sentiment analysis, and semantic tagging to identify racist and radicalized Tumblr posts based on their narrative intent. In this scheme, a cascaded ensemble classification model has been presented for topic classification, and intent classification tasks. In another study, Alfina et al. [9] constructed a novel dataset for several aspects of hate speech, including religion, race, ethnicity, and gender. The predictive performance of conventional linguistic features including word bigram, word unigram, character trigram, character four-gram, and negative sentiment aspects have been considered. In the empirical analysis, supervised learning algorithms, such as, Naïve Bayes, support vector machines, Bayesian logistic regression, and random forest algorithm have been evaluated. The empirical results indicated that word n-gram features can outperform character n-gram models. Recently, Oriola and Kotze [10] evaluated the predictive performance of word n-gram features, character n-gram features, syntactic features, and negative sentiment-based features for identification of offensive and hate speech in South African tweets.

Deep neural networks are being employed to identify offensive and hateful speech on social media platforms as well. For instance, Badjatiya et al. [11] presented a deep learning-based approach for hateful speech identification. In this scheme, the predictive performance of three neural language models (i.e., random embedding, fastText, and GloVe) have been evaluated in conjunction with two deep neural network architectures, namely, convolutional neural networks and long short-term memory architecture.

In another study, Kapil et al. [12] examined the conventional text representation schemes and neural language models for hate speech identification. In this regard, word2vec, GloVe, fastText, and one-hot encoding have been considered. The performance of 13 different deep neural architectures based on convolutional neural networks, long short-term memory, bidirectional long short-term memory, and character-convolutional neural networks have been analyzed. The experiment analysis indicated that GloVe word embedding scheme in conjunction with bidirectional long short-term memory and character-convolutional neural network yields promising results for hate speech

recognition task. In a similar way, Kapil and Ekbal [13] introduced a novel hybrid deep neural network architecture based on convolutional neural network, gated recurrent unit and multi-task learning for hate speech detection. In a similar way, Lee and Lee [14] presented a deep learning-based scheme for bias and hate speech identification. In this scheme, multi-channel convolutional neural network architecture with the attention mechanism has been presented. Recently, Zhao and Tao [15] presented a deep learning-based scheme for offensive language identification. In this scheme, XML-RoBERTa pre-training model has been utilized.

The identification of offensive and hateful speech on social media can be regarded as a text classification task. Classification of texts into pre-defined categories or classes is a critical subfield of text mining. Web page classification [16], sentiment analysis [17-22], and text genre classification [23] have all used text mining to great effect. It is critical to use an appropriate word embedding scheme and a deep learning architecture to identify hateful speech. In this context, these experiments will evaluate the usefulness of conventional word embedding and deep learning architectures.

3. METHODOLOGY

The section contains dataset used to examine the offensive and hateful language identification, as well as description of the text embedding methods and deep learning architectures that were utilized in the empirical study.

3.1. Dataset

In the empirical analysis, the hate speech lexicon obtained from hatebase.org has been utilized [5]. In this dataset, user-provided words and phrases have been utilized to annotate the text documents as hate speech. The dataset has been obtained via Twitter API. After pre-processing stages, a randomly sampled collection of 25,000 tweets have been included in the dataset. The “hate” label is assigned to tweets that contain hateful text, the “offensive” label is assigned to tweets that contain only offensive text, and the tweets that are neither hateful nor offensive are annotated as “okay” in the dataset.

3.2. Neural language models

The conventional representation scheme in machine learning-based natural language processing (NLP) tasks is to represent text documents via bag-of-words (BOW). BOW has two significant drawbacks. In BOW, the order of the words does not matter. In addition, the total number of words in the text collection corresponds to the vector length, which involves high dimensionality.

In deep learning-based NLP tasks, neural language models have been frequently employed, which enable the representation of data with fewer dimensions [24]. We have examined three approaches for embedding words (namely, word2vec, GloVe, and fastText) in the empirical analysis.

The Word2vec model is a three-layer artificial neural network-based technique for word embedding [25]. It is composed of three layers: an input one, an output one, and a hidden one. Its goal is to develop an understanding of word embedding by calculating the probability that a given word will be surrounded by other words. The model is based on two fundamental architectures: the skip-gram (SG) and the continuous-bag-of-words (CBOW). The CBOW architecture defines the target word by analyzing its content; by contrast, the SG architecture takes the target word as input and predicts the words that surround it. The CBOW architecture is capable of handling small amounts of data efficiently. On the other hand, the SG architecture outperforms the CBOW architecture when dealing with large data sets.

The GloVe is a word2vec-based model that was developed to learn word embeddings from text documents efficiently. The model incorporates local contextual learning from the word2vec model with global matrix factorization [26]. When calculating the error function of the model, the probability ratios of words are also considered. Words that appear close together in a text document and are more likely to be seen together are more important than other words during the learning process [25].

The fastText model is another efficient technique for obtaining word embeddings from text documents. Each word is represented by dividing the character into n-grams in this model. We generate word vectors for each n-gram in the training set. The fastText model provides a more efficient word embedding scheme for morphologically rich languages and rare words [27].

3.3. Deep learning architectures

Deep learning architectures have recently gained considerable research attention in a variety of fields, including text classification. In the empirical analysis, convolutional neural networks, recurrent neural networks, long short-term memory, gated recurrent unit, bidirectional long short-term memory and bidirectional gated recurrent units have all been examined.

Convolutional neural networks (CNNs) are a subclass of deep neural networks that process data using a grid-based topology [28]. Convolution, a specialized type of mathematical operation, was used in one or more convolutional layers of the CNN instead of general matrix multiplication, as is the case with conventional neural network architectures. Three layers comprise the architecture of the CNN: input, output, and hidden. The architecture's hidden layers function as substitutes for several layers, including convolutional layers, pooling layers, fully connected layers, and normalization layers.

A recurrent neural network (RNN) is a type of deep network architecture designed for the purpose of processing sequential data. The nerve cells' connections form a directed graph in this architecture. It is possible to process inputs of any length using the RNN architecture. Historic data is also considered during the calculation [24]. The vanishing gradient problem frequently manifests itself in RNN architectures. Long-term dependencies are difficult to model in RNN architectures because the multiplicative gradient varies with the number of layers [27].

Long short-term memory units (LSTM) are an iterative neural network architecture that circumvents the vanishing gradient problem that plagues traditional RNN architectures. With a finite number of time steps, LSTM enables error propagation backwards. A fundamental LSTM unit consists of a cell and three fundamental gates: an entry gate, an exit gate, and a forget gate. It is determined which data should be protected and when the units will be accessed based on the condition of the doors [27].

The gated repeating unit (GRU) architecture is a recurrent neural network architecture that is less complex than the LSTM and produces comparable experimental results [27, 29]. A basic GRU architecture consists of two ports: one for reset and one for update.

The bidirectional architectures (namely, Bi-LSTM and Bi-GRU) utilizes hidden layers in both directions to capture both potential and previous contexts. As a result, the network learns more efficiently, and the transfer of temporal knowledge is increased in both directions [27].

4. EXPERIMENTAL PROCEDURE AND RESULTS

This section describes the experimental design, and the empirical findings.

4.1. Experimental procedure

The deep learning architectures used in the empirical analysis were developed and trained using TensorFlow and Keras. We used a hyper parameter searching algorithm to optimize the predictive performance of each deep learning model. This was accomplished using hyper parameter optimization based on Bayesian optimization via the Gaussian process. Eighty percent of the data in the corpus was used as a training set, while the remaining twenty percent was used as a testing set. We considered the word2vec and fastText schemes, as well as the continuous skip-gram and CBOW schemes, with varying vector sizes (200 and 300) and projection layer dimensions (100 and 200). To evaluate the performance of classification algorithms, two distinct evaluation metrics, namely, classification accuracy and the F-measure have been considered.

Table 1. Classification accuracy values obtained by compared word embedding schemes and deep neural networks

Word embedding	Vector size	Layer size	CNN	RNN	LSTM	GRU	Bi-LSTM	Bi-GRU
word2vec (Skip-gram)	200	100	79,48	82,38	83,80	85,31	86,80	88,74
word2vec (Skip-gram)	200	200	80,28	83,02	84,28	85,54	87,25	89,13
word2vec (Skip-gram)	300	100	81,38	83,24	84,43	86,23	87,45	89,74
word2vec (Skip-gram)	300	300	81,95	83,36	84,98	86,36	88,26	92,30
word2vec (CBOW)	200	100	79,99	82,44	84,04	85,38	86,81	88,74

word2vec (CBOW)	200	200	80,47	83,03	84,29	85,73	87,33	89,20
word2vec (CBOW)	300	100	81,38	83,25	84,50	86,24	87,71	89,88
word2vec (CBOW)	300	300	82,10	83,49	85,04	86,37	88,35	92,33
fastText (Skip-gram)	200	100	80,03	82,82	84,05	85,39	86,93	88,79
fastText (Skip-gram)	200	200	80,72	83,05	84,37	85,80	87,35	89,27
fastText (Skip-gram)	300	100	81,74	83,31	84,77	86,26	87,95	90,67
fastText (Skip-gram)	300	300	82,12	83,57	85,10	86,50	88,37	92,66
fastText (CBOW)	200	100	80,03	82,85	84,06	85,42	87,06	88,83
fastText (CBOW)	200	200	81,11	83,10	84,39	85,91	87,36	89,49
fastText (CBOW)	300	100	81,74	83,33	84,88	86,32	87,97	91,49
fastText (CBOW)	300	300	82,22	83,62	85,23	86,54	88,40	93,02
GloVe	200	100	80,11	82,99	84,26	85,50	87,17	88,84
GloVe	200	200	81,13	83,23	84,40	86,00	87,36	89,65
GloVe	300	100	81,85	83,36	84,94	86,34	88,10	92,15
GloVe	300	300	82,25	83,77	85,24	86,62	88,65	93,84

4.2. Experimental procedure

In Tables 1 and 2, the classification accuracy and F-measure values obtained by the compared word embedding schemes and deep neural architectures have been presented, respectively. Regarding the word embedding schemes considered in the empirical results, the highest results have been generally achieved by GloVe word embedding scheme. The second highest predictive performance has been generally achieved by fastText (CBOW) model, which is followed by fastText (skip-gram) model. The lowest predictive performances have been generally achieved by word2vec neural language model. Regarding the different vector sizes considered in the empirical analysis, vector size of 300 yields higher predictive performance compared to vector size of 200. The third factor of the empirical analysis is to examine the dimension of the projection layer. In this regard, we have considered projection layer size of 100, 200, and 300. The highest predictive performance among these compared configurations has been achieved by the projection layer size of 300. Regarding the deep neural network architectures employed in the analysis, bidirectional architectures outperform the conventional architectures. The highest predictive performance among all the compared has been achieved by bi-directional gated recurrent neural network architecture in conjunction with GloVe word embedding model. In Figures 1 and 2, the main effects plot for accuracy and F-measure values have been presented to summarize the findings of the empirical analysis.

Table 2. F-measure values obtained by compared word embedding schemes and deep neural networks

Word embedding	**Vector size**	**Layer size**	**CNN**	**RNN**	**LSTM**	**GRU**	**Bi-LSTM**	**Bi-GRU**
word2vec (Skip-gram)	200	100	0,80	0,83	0,84	0,86	0,87	0,89
word2vec (Skip-gram)	200	200	0,81	0,84	0,85	0,86	0,88	0,90
word2vec (Skip-gram)	300	100	0,82	0,84	0,85	0,87	0,88	0,90
word2vec (Skip-gram)	300	300	0,83	0,84	0,86	0,87	0,89	0,93
word2vec (CBOW)	200	100	0,81	0,83	0,85	0,86	0,87	0,89
word2vec (CBOW)	200	200	0,81	0,84	0,85	0,86	0,88	0,90
word2vec (CBOW)	300	100	0,82	0,84	0,85	0,87	0,88	0,91
word2vec (CBOW)	300	300	0,83	0,84	0,86	0,87	0,89	0,93
fastText (Skip-gram)	200	100	0,81	0,83	0,85	0,86	0,88	0,89
fastText (Skip-gram)	200	200	0,81	0,84	0,85	0,86	0,88	0,90
fastText (Skip-gram)	300	100	0,82	0,84	0,85	0,87	0,89	0,91
fastText (Skip-gram)	300	300	0,83	0,84	0,86	0,87	0,89	0,93
fastText (CBOW)	200	100	0,81	0,83	0,85	0,86	0,88	0,89
fastText (CBOW)	200	200	0,82	0,84	0,85	0,87	0,88	0,90
fastText (CBOW)	300	100	0,82	0,84	0,85	0,87	0,89	0,92
fastText (CBOW)	300	300	0,83	0,84	0,86	0,87	0,89	0,94
GloVe	200	100	0,81	0,84	0,85	0,86	0,88	0,89
GloVe	200	200	0,82	0,84	0,85	0,87	0,88	0,90
GloVe	300	100	0,82	0,84	0,86	0,87	0,89	0,93
GloVe	300	300	0,83	0,84	0,86	0,87	0,89	0,95

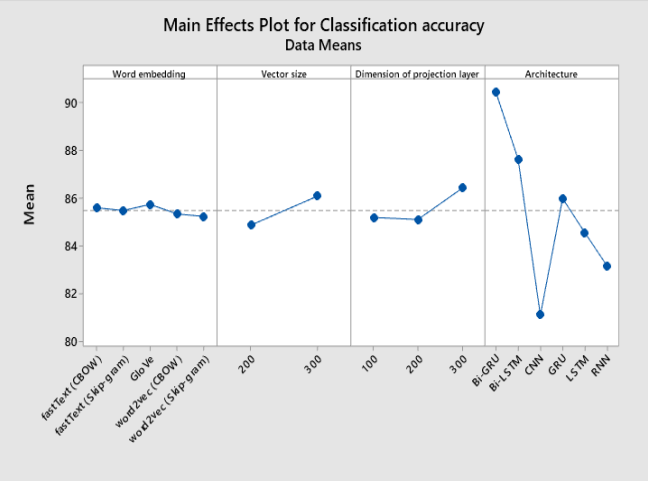

Fig. 1. Main effects plot for classification accuracy values

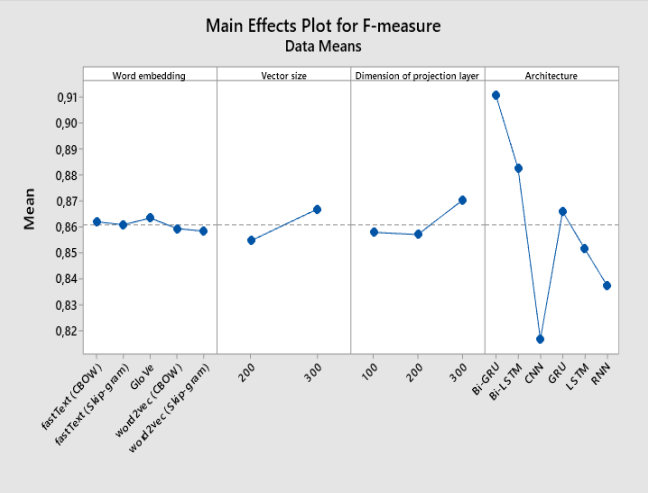

Fig. 2. Main effects plot for classification F-measure values

CONCLUSION

While the Internet and mobile devices have greatly simplified access to social media platforms, which facilitates the exchange of ideas, opinions, and feelings, such platforms are also prone to abuse, and offensive and hateful language can be easily shared on social media. Unchecked growth of hate has the potential to cause significant harm to our society, as well as to marginalized groups and individuals. The practice of manually and algorithmically reviewing content to remove offensive language and content has become widespread on social media. Manual methods are destined to become increasingly inefficient as the number of users and their posts on social media platforms continue to grow. As a result, identifying offensive and hateful language on social media platforms is a critical task for natural language processing. We examine the predictive performance of deep neural networks in identifying offensive and hateful language in this article. In the empirical analysis, three distinct word embedding schemes were compared: word2vec, GloVe, and fastText. The following algorithms have been examined: convolutional neural networks, recurrent neural networks, long short-term memory, gated recurrent units, bidirectional long short-term memory, and bidirectional gated recurrent units. The empirical evidence suggests that bidirectional gated recurrent units can perform well on this task. The highest predictive performance among all the compared has been achieved by bi-directional gated recurrent neural network architecture in conjunction with GloVe word embedding model.

References

[1] A. H. Wani, N. S. Molvi, and S. I. Ashraf, "Detection of hate and offensive speech in text," in Intelligent Human Computer Interaction, Cham: Springer International Publishing, 2020, pp. 87–93.

[2] G. Kovács, P. Alonso, and R. Saini, "Challenges of hate speech detection in social media: Data scarcity, and leveraging external resources," SN Computer Science, vol. 2, no. 2, 2021.

[3] S. Ullmann and M. Tomalin, "Quarantining online hate speech: technical and ethical perspectives," Ethics Inf. Technol., vol. 22, no. 1, pp. 69–80, 2020.

[4] C. Nobata, J. Tetreault, A. Thomas, Y. Mehdad, and Y. Chang, "Abusive language detection in online user content," in Proceedings of the 25th International Conference on World Wide Web - WWW '16, 2016.

[5] T. Davidson, D. Warmsley, M. Macy, and I. Weber, "Automated hate speech detection and the problem of offensive language," arXiv [cs.CL], 2017.

[6] S. Sharma, S. Agrawal, and M. Shrivastava, "Degree based Classification of Harmful Speech using Twitter Data," arXiv [cs.CL], 2018.

[7] Z. Waseem and D. Hovy, "Hateful symbols or hateful people? Predictive features for hate speech detection on twitter," in Proceedings of the NAACL Student Research Workshop, 2016.

[8] S. Agarwal and A. Sureka, "Characterizing linguistic attributes for automatic classification of intent based racist/radicalized posts on Tumblr micro-blogging website," arXiv [cs.IR], 2017.

[9] I. Alfina, R. Mulia, M. I. Fanany, and Y. Ekanata, "Hate speech detection in the Indonesian language: A dataset and preliminary study," in 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2017.

[10] O. Oriola and E. Kotze, "Evaluating machine learning techniques for detecting offensive and hate speech in south African tweets," IEEE Access, vol. 8, pp. 21496–21509, 2020.

[11] P. Badjatiya, S. Gupta, M. Gupta, and V. Varma, "Deep learning for hate speech detection in tweets," arXiv [cs.CL], 2017.

[12] P. Kapil, A. Ekbal, and D. Das, "Investigating deep learning approaches for hate speech detection in social media," arXiv [cs.CL], 2020.

[13] P. Kapil and A. Ekbal, "A deep neural network based multi-task learning approach to hate speech detection," Knowl. Based Syst., vol. 210, no. 106458, p. 106458, 2020.

[14] W. Lee and H. Lee, "Bias & hate speech detection using deep learning: Multi-channel CNN modeling with attention," J. Korean Inst. Inf. Commun. Eng., vol. 24, no. 12, pp. 1595–1603, 2020.

[15] Y. Zhao and X. Tao, "ZYJ123@DravidianLangTech-EACL2021: Offensive Language Identification based on XLM-RoBERTa with DPCNN," in Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages, 2021, pp. 216–221.

[16] A. Onan, "Classifier and feature set ensembles for web page classification," J. Inf. Sci., vol. 42, no. 2, pp. 150–165, 2016.

[17] M. A. Toçoğlu and A. Onan, "Sentiment analysis on students' evaluation of higher educational institutions," in Advances in Intelligent Systems and Computing, Cham: Springer International Publishing, 2021, pp. 1693–1700.

[18] A. Onan, "Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks," Concurr. Comput., 2020.

[19] A. Onan, "Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach," Comput. Appl. Eng. Educ., no. cae.22253, 2020.

[20] A. Onan, "Sentiment analysis on twitter based on ensemble of psychological and linguistic feature sets," Balk. J. Electr. Comput. Eng., pp. 1–9, 2018.

[21] A. Onan and S. Korukoğlu, "A feature selection model based on genetic rank aggregation for text sentiment classification," J. Inf. Sci., vol. 43, no. 1, pp. 25–38, 2017.

[22] A. Onan, S. Korukoğlu, and H. Bulut, "A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification," Expert Syst. Appl., vol. 62, pp. 1–16, 2016.

[23] A. Onan, "An ensemble scheme based on language function analysis and feature engineering for text genre classification," J. Inf. Sci., vol. 44, no. 1, pp. 28–47, 2018.

[24] A. Onan, "Mining opinions from instructor evaluation reviews: A deep learning approach," Comput. Appl. Eng. Educ., vol. 28, no. 1, pp. 117–138, 2020.

[25] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," arXiv [cs.CL], 2013.

[26] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014.

[27] S. Sukheja, S. Chopra, and M. Vijayalakshmi, "Sentiment Analysis using Deep Learning - A survey," in 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), 2020.

[28] D. Ciresan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2012.

[29] G. Jain, M. Sharma, and B. Agarwal, "Spam detection in social media using convolutional and long short term memory neural network," Ann. Math. Artif. Intell., vol. 85, no. 1, pp. 21–44, 2019.

Factory Environment Safety via Edge Computing Technology

Yang-Han LEE[a], Cheng-Wen CHEN [b]1, Yi-Lun CHEN[c], Chen-Shiuan WANG[d]

[a] *Department of Electrical and Computer Engineering, Tamkang University, Taiwan, R.O.C., yhlee@ee.tku.edu.tw*
[b] *Department of Electrical and Computer Engineering, Tamkang University, Taiwan, R.O.C., 809440042@gms.tku.edu.tw*
[c] *Department of Electrical and Computer Engineering, Tamkang University, Taiwan, R.O.C., alanchen0318@gmail.com*
[d] *Department of Electrical and Computer Engineering, Tamkang University, Taiwan, R.O.C., bruce61002@gmail.com*

Abstract

With the development of the automation industry, the number of worker in the factory is reduced. Therefore, for each worker it is not easy to detect and put out the fire immediately. We design a system with Bluetooth sensors and Bluetooth broadcast protocol. This system can sense environmental data and transmit the data to the edge computer through Bluetooth broadcast. The edge computer will scan the surrounding Bluetooth signals to identify the Bluetooth sensor broadcast signals that contain environmental data. After the edge computer obtains the environmental data, it will use the algorithms to calculate environmental changes and make the intelligent decision. If the edge computer find out the environment to be in danger, it will send warning message to the users. The edge computer will notify the users. And it will stop the PLC equipment. So the operation of the production line will be stopped to reduce the loss of the factory.

Keywords: Environment Safety, IoT, Edge computing, Bluetooth, PLC

1. INTRODUCTION

The fire safety of the factory is extremely important. Even with strict regulations and sufficient fire protection awareness, fire accidents will still occur. The negligence of personnel to pay attention to the details of the factory is one of the reasons for accidental fire alarms in the factory, such as the temperature in the environment. It will affect the safety of the factory's storage. The humidity level will affect the incidence of electrostatic discharge bombardment; however, workers cannot always monitor environmental changes and constantly inspect the equipment. Bluetooth sensors and edge computers capable of intelligent judgment assist workers in monitoring the factory environment.

This paper designs a smart firefighting system. With the wireless transmission of Bluetooth sensors and easy deployment, the edge computer on the production line makes intelligent judgments of the environment. When the environment changes too fast, a message is sent to notify users to pay attention. When environmental data exceeds the warning value, an alarm is issued to notify the user. When necessary, the edge computer will actively control the Programmable Logic Controller (PLC) of the production line to perform related actions to achieve the intelligent automatic fire protection system.

* Corresponding author.

Reference [1] uses Arduino Mega and Raspberry Pi 3B+ to design a fire alarm system. Arduino uses multiple optical sensors to detect flames and chemical smoke to determine the location of abnormal environments, and use an automatically controlled lens module to observe abnormalities. Direction, finally connect the Arduino IP to the Router using Wi-Fi via ESP-01. Reference [2] designed an automatic fire monitoring system for the textile industry, reference [3] [4] designed a fire alarm and fire control system, reference [5] is for home fire monitoring, reference [6] is aimed at For the monitoring of hazardous materials, reference [7] is a simulation of fire in household appliances factories, reference [8] is a wireless fire monitoring system, reference [9] is a fire extinguishing system using Internet of Things technology, reference [10]][11][12] uses wireless network to realize, reference [13][14][15][16] uses Arduino to realize related system.

2. RESEARCH DESIGN

2.1. System architecture diagram

This research is divided into three parts: the Bluetooth sensor responsible for collecting environmental data, the edge computer for data analysis, and the PLC responsible for dealing with abnormal conditions on the production Line, and it is combined with Line Notify to achieve the function of reporting users, as shown in Figure 1.

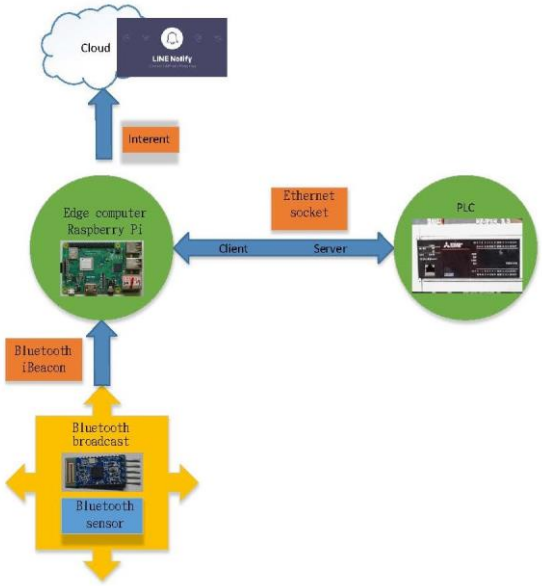

Fig. 1. System architecture diagram

2.2. System flow chart

The Bluetooth sensor is responsible for collecting the response data and sending the data to the edge computer using Bluetooth broadcasting. When the edge computer receives the environmental data, it will store the data and make a judgment. When the data is judged to be abnormal, it will report to the user through Line Notify Perform related processing with the control PLC, and the flowchart is shown in Figure 2.

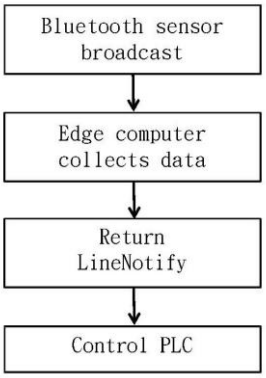

Fig. 2. System flow chart

2.3. Bluetooth sensor

Bluetooth sensors are different from traditional fire sensors that use a single sensing method such as optical and thermal. Bluetooth sensors use multiple sensor modules, as shown in Figure 3, which continuously senses a variety of environmental data; And different from the traditional fire sensor that transmits the fire alarm to the fire protection switchboard by wire when it detects a fire, the Bluetooth sensor in this research uses the Bluetooth signal as the transmission method, which achieves the characteristics of easy deployment and diversification of the equipment.

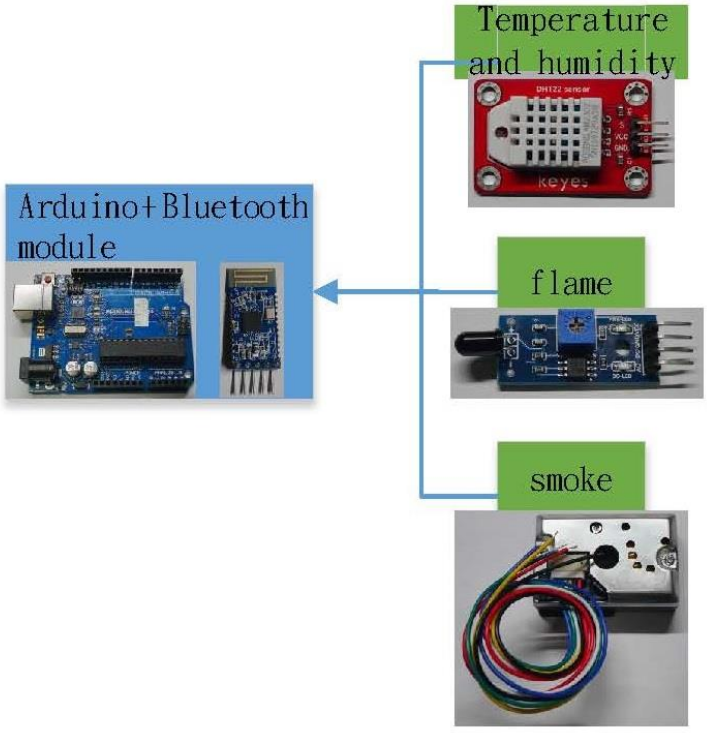

Fig. 3. Bluetooth sensor

2.4. Edge computer

The edge computer is different from the traditional fire protection switchboard by connecting the sensor by wire. The edge computer can scan and collect the Bluetooth sensor through its built-in Bluetooth device to send a

Bluetooth signal containing environmental data, and perform data calculation and judgment, and Perform related actions according to the judgment result, such as reminding the user that the environment changes too fast and requiring special attention, so as to remind the user to take preventive actions, and use the lens module to send real-time images to the remote user, as shown in Figure 4.

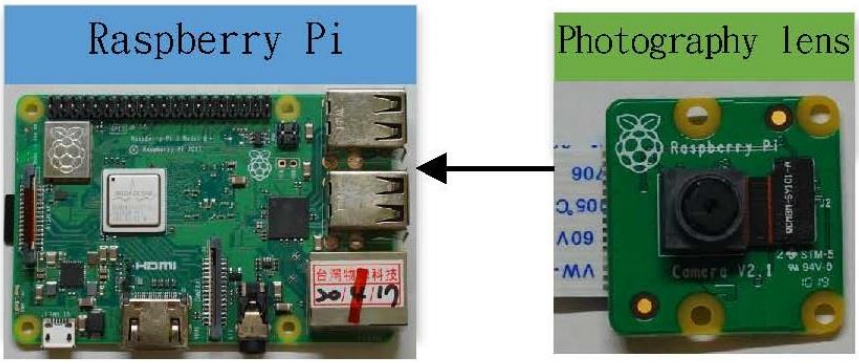

Fig. 4. Edge computer

2.5. Programmable logic controller

With the development of Industry 4.0, factories have gradually become automated, and PLC has become one of the best solutions for automated factories with flexible and stable systems, used to control production line machines. When a fire occurs, the PLC connected to the edge computer can be controlled by the edge computer to minimize the equipment loss of the production line.

3. RESEARCH METHODS

3.1. Bluetooth sensor

The Bluetooth sensor uses the Arduino UNO R3 development version, with DHT-22 temperature and humidity sensor, Sharp GP2Y1014AU smoke sensor, KY-23 flame sensor, and JDY-18 Bluetooth module to collect various data in the environment. And send the data to the edge computer by way of Bluetooth broadcast, as shown in Figure 5.

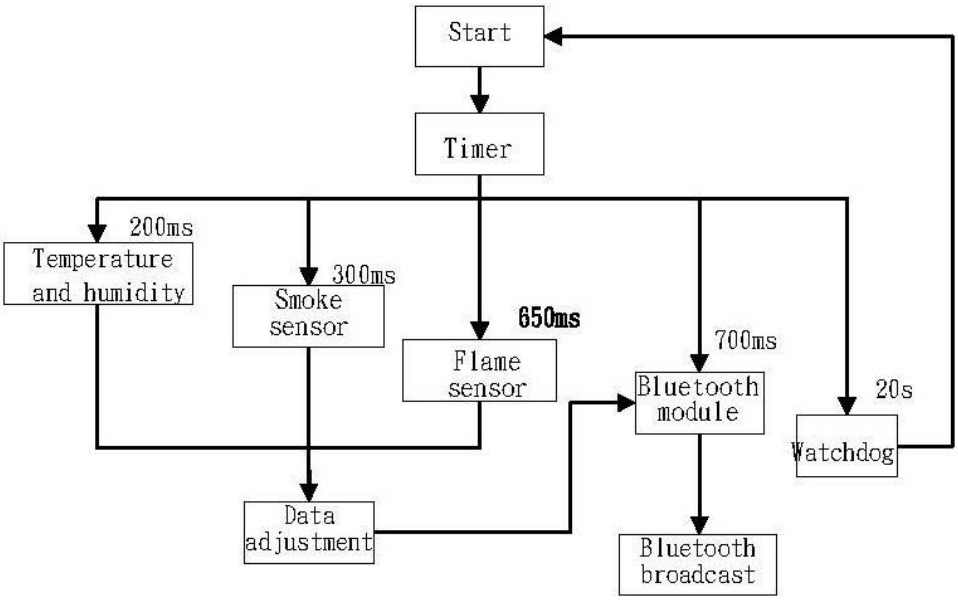

Fig. 5. Bluetooth sensor flow chart

Use the timer on the Arduino UNO to read the temperature and humidity sensor at 200 milliseconds per second, read the smoke sensor at 300 milliseconds, read the flame sensor at 650 milliseconds, and update the Bluetooth module data at 700 milliseconds. And start the watchdog every 20 seconds to confirm the device status; the Bluetooth broadcast will continue to be sent, but the broadcast will stop when the Bluetooth data is updated.

The development board uses Arduino UNO R3. Arduino UNO is an open source control board developed by Arduino.cc using Microchip ATmega328P microcontroller. This development version has 14 digital pins and 6 analog pins. It supports SPI, UART and IIC communication protocols. It can be connected to a computer with its own B-type USB interfaces. It can be used to write programs with the open source software Arduino IDE to achieve soft the purpose of rapid hardware development.

The temperature and humidity sensor uses DHT-22, which is an Aosong AM2302 digital temperature and humidity sensor that uses a digital output signal that has been calibrated before leaving the factory. The sensor includes an NTC temperature measuring element and a capacitive humidity sensing element. The calibration data is stored in the OTP memory in a programmed form and connected to a high-performance 8-bit single-chip microcomputer. Therefore, the sensor has the advantages of fast response, easy development, small size, and low power consumption.

The smoke sensor uses SHARP GP2Y1014AU, which uses optical methods to detect air particles. Its internal design is equipped with diagonal infrared diodes and photoelectric crystals. It is detected by the reflected light of suspended particles in the air and can measure tiny particles above 0.8 microns. Particles, and the sensor have the advantages of low power consumption and small size.

The flame sensor uses KY-23, which is suitable for all kinds of flames. It uses the flame to have strong infrared characteristics, uses an infrared receiver tube to sense the flame, and then converts the sensed flame infrared intensity into a high and low level signal.

The Bluetooth module uses JDY-18, which is based on the Bluetooth 4.2 communication protocol and is compatible with Bluetooth 5.0. The working frequency band is 2.4GHz. The user can use AT commands to set the device name, transmit power, working mode, pairing password and other functions. The working modes include master-slave pairing mode, iBeacon broadcast and probe mode.

3.2. Bluetooth broadcast technology Beacon

Bluetooth Beacon is a broadcast technology based on the Bluetooth Low Energy (BLE) protocol. It works as a device that already has Bluetooth low energy communication capabilities and uses Bluetooth low energy technology to send its own unique data packets to surrounding devices. The data packet contains the device-specific service UUID (Service Universally Unique Identifier), short data, and RSSI, which is quite short. Beacon will broadcast the

data packet to the surroundings at regular intervals, and use the supported Bluetooth device to receive the Beacon data packet, which is often used for accurate indoor positioning or for product promotion with mobile APP.

Apple's iBeacon is the earliest company on the market to use Beacon technology. The iBeacon packet contains UUID, RSSI and other formats. The short data is divided into 16bits Major and Minor; the short data was originally used to push product codes and match RSSI and mobile APP programs to provide product information, etc. However, in this study, Major and Minor are used as the tools for sensor data transmission. The raw data of each sensor module occupies too much space and cannot be put into the Major and Minor completely. In order to smoothly transmit the temperature, humidity, smoke, and flame data through the iBeacon protocol, the following changes the data of each sensor. The temperature and humidity data of the temperature and humidity sensor use 32bits floating-point numbers. Because the accuracy of the sensor reaches one decimal place and the range is 0 to 100, both data are magnified ten times and converted to Integer data, the occupied space becomes 10bits. After the temperature and humidity data occupy 10bits each for Major and Minor, there is still 6bits space left to provide smoke sensors and flame sensors. The range of values detected by the smoke sensor is an integer of 0~3000 (12bits). In this study, it will be proportionally mapped to 0~63, occupying the remaining 6bits space of Major. The range of values detected by the flame sensor is an integer of 0~1023 (10bits). In this study, it is finally mapped to 0~63, occupying the remaining 6bits space of the Minor, as shown in Figure 6.

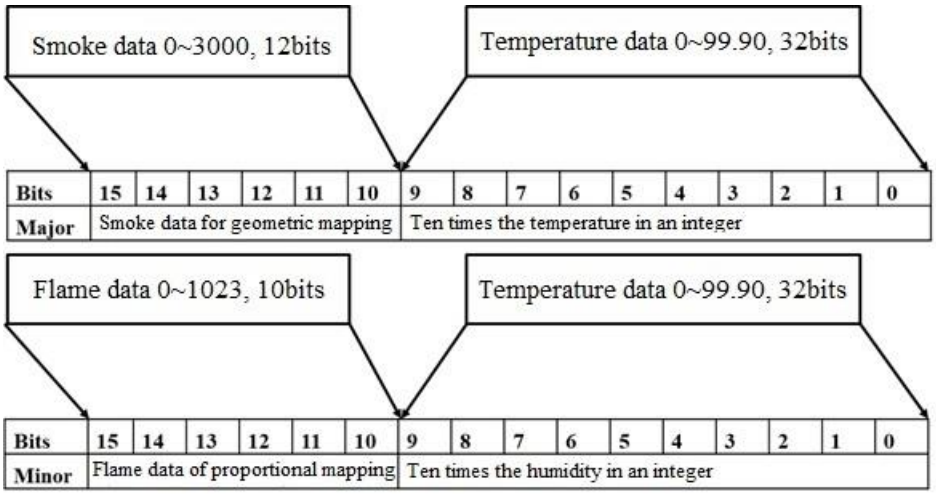

Fig. 6. iBeacon data space allocation

3.3. Raspberry Pi Edge Computer

This research uses Raspberry Pi 3B+ with Pi Camera V2 lens module as the edge computer, and uses three important tools to achieve the goal, namely Bluepy, SQLite and Line Notify. Raspberry Pi 3B+ is a single board computer using Linux system, with ARM Cortex-A53 processor, 1GB memory, WiFi, Bluetooth 4.2BLE and Ethernet interface. The Pi Camera V2 lens module has 8 million pixels and is connected to the Raspberry Pi through the CSI interface, which is used in this research to capture flame images and report to the user. The main workflow of the edge computer has three parts. The first is to scan and collect Bluetooth sensor data. The second is to perform an algorithm on the collected data to determine whether the environmental data is abnormal. The third is when the data judgment result is abnormal. To inform users or control related equipment.

Bluepy is open source software developed by GitHub IanHarvey, which realizes the use of Python interface to control BLE devices under Linux environment. The edge computer will use Bluepy to control the Bluetooth low energy device on the Raspberry Pi, scan all Bluetooth packets in the surrounding environment, and save the packets with the sensor-specific UUID tag to obtain the data of all the sensor modules.

SQLite is a portable relational database. It has the characteristics of serverless and cross-platform. It can store data on any client, providing a portable and easy-to-use simple database. After the edge computer has parsed the Bluetooth broadcast packet from the Bluetooth sensor, it will obtain the sensor's Bluetooth address and the data of each sensor. The edge computer will use the Bluetooth address as the name to create a database, and according to time, smoke sensing the sequence of the sensor value, temperature, flame sensor value, and humidity is stored in the database.

Line Notify is a service provided by Line. After users apply for this function, Line will issue an access token to provide users. Users can write automatic scripts and use Line Notify to send any message. Anything that has the same access token here In the Line Notify group, you will receive messages or photos of automatic scripts. In this study, Line Notify achieves the function of notifying users. When Bluepy scans and collects the data of the Bluetooth sensor and saves the relevant values into SQLite, it will judge the data. If the judgment result is abnormal, the abnormal value will be passed through Line Notify reports to users.

There are two types of abnormal report values. The first is to determine the normal distribution in statistics, and calculate the normal distribution of the latest 100 Bluetooth sensor data. If the scanned Bluetooth sensor data exceeds this The subnormal distribution will be judged as the environment changes too fast; to calculate the normal distribution, the average value (1) and the standard deviation (2) are required, and the two can be expressed as the following two formulas , the normal distribution is 3 standard deviations from the mean, and the coverage distribution ratio is 99.7%.

$$\text{avg} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

(1)

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - avg)^2}$$

(2)

The second is to set the abnormal value limit in response to environmental conditions. If the abnormal value limit is exceeded, it will be directly judged as an environmental abnormality: Temperature: According to the "High Temperature Work Rest Time Standard", the maximum comprehensive temperature thermal index value for light work is 33°C, but Set at 40°C due to the experimental environment. Humidity: The temperature rises and the humidity drops during a fire, so the humidity alarm is set at 30% RH lower than the normal environment. Flame: Due to the characteristics of the sensor, indoor lighting will cause some errors. To avoid misjudgment, set the value to 15 levels. Smoke: According to the corresponding data in the data sheet, poor air quality corresponds to a value of 30.

3.4. Programmable logic controller

PLC is a kind of microprocessor digital electronic device that can be controlled by logic programming. It is a digital controller for automatic control. With the continuous improvement of factory automation, the use of PLC is becoming more and more popular. This experiment uses PLC to simulate a production line with fire extinguishing devices. By constantly communicating with the edge computer to obtain work instructions, when the edge computer knows that the Bluetooth sensor detects the flame, it will issue a fire extinguishing action to the PLC through the socket interface. The PLC will also report that the fire is being extinguished; until the Bluetooth sensor no longer detects the flame, the edge computer will issue an end fire extinguishing action, and the PLC will also report a shutdown until the PLC is restored to normal operation.

4. ENVIRONMENTAL DESIGN SIMULATION RESULTS

4.1. Abnormal temperature changes

Put the Bluetooth sensor in the transparent acrylic box as shown in Figure 7. Start and leave it for more than 30 minutes, let the edge computer collect data and obtain the normal temperature distribution; when the edge computer obtains the normal distribution, it will be placed in the acrylic The hair dryer at the opening of the force box starts to heat the inside of the acrylic box. After the hair dryer starts to heat, the temperature exceeds the normal distribution range of the previous environmental statistics of the edge computer, and a Line Notify message is sent to warn the user that the ambient temperature changes too fast. When the temperature exceeds forty degrees, the user will be immediately reported to the user that the ambient temperature is too high. , The data record of the experiment process is shown in Figure 8.

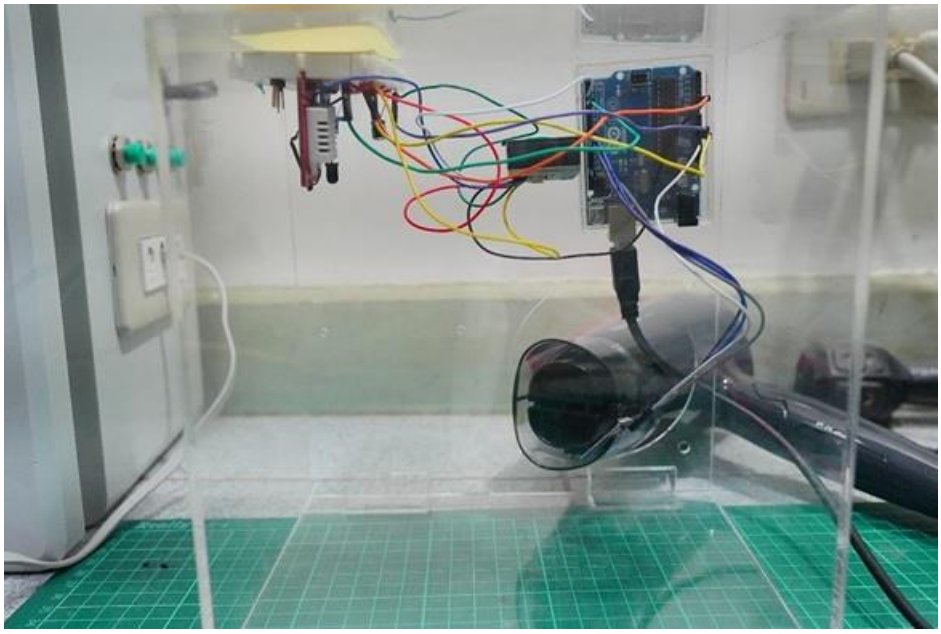

Fig. 7. Field of abnormal temperature change

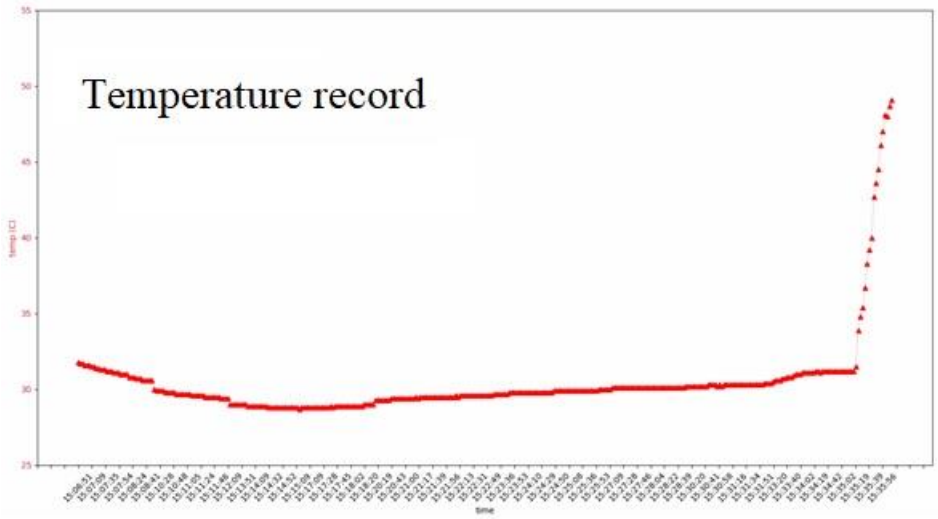

Fig. 8. Abnormal change simulation data record

4.2. The smoke and flame sensor exceeds the alarm value

Place the Bluetooth sensor in a transparent acrylic box to simulate the field as shown in Figure 9. First click on the incense to simulate smoke when a fire occurs. The edge computer immediately sends a Line Notify message to inform the user that there is smoke in the environment; detection; when the smoke of the incense stick is reached, the candle is lit to simulate the existence of a flame. After the edge computer judges that there is a flame, it immediately activates the lens module to shoot the scene and report the user's scene status through Line Notify. The data record of the experiment is shown in Figure 10.

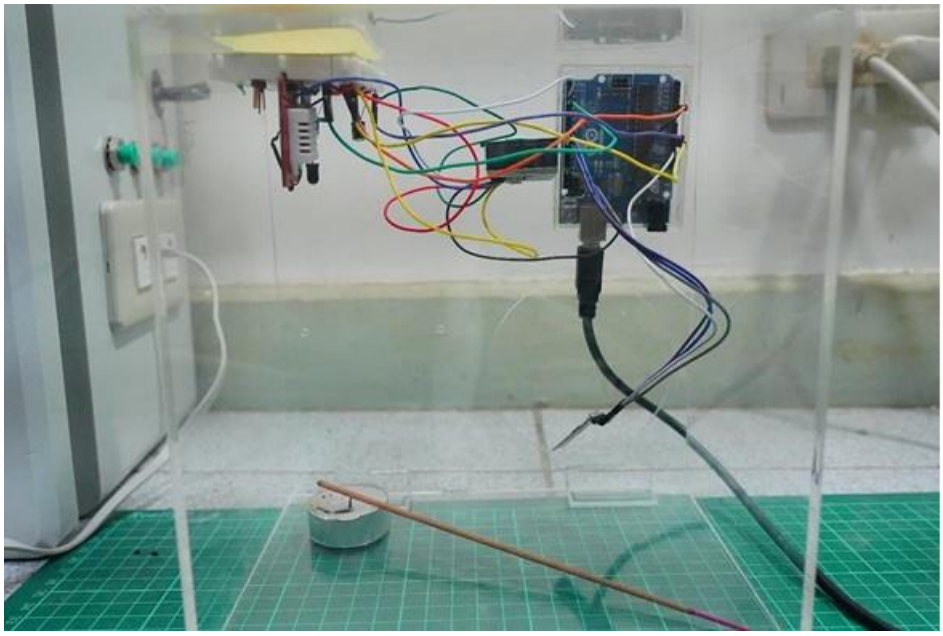

Fig. 9. Smoke and flame field

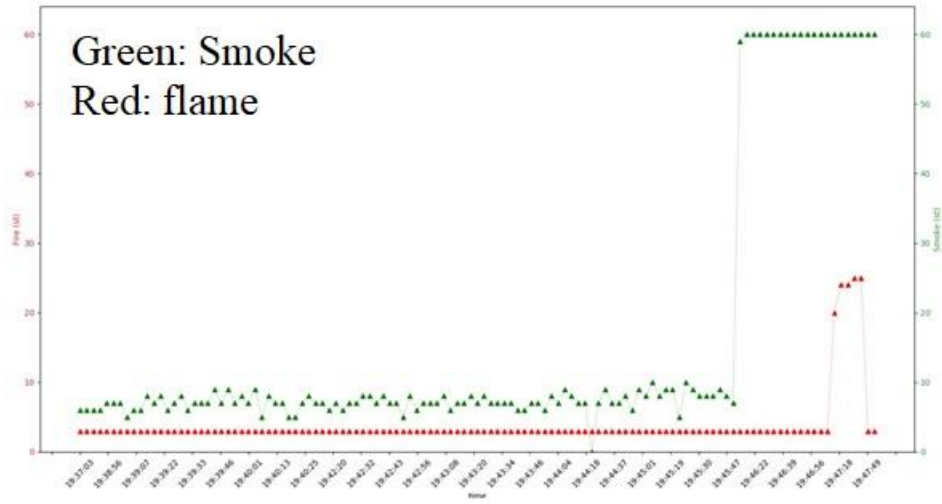

Fig. 10. Smoke and flame sensing records

4.3. Edge computer control PLC fire extinguishing

As shown in Figure 11(a), the Bluetooth sensor and candle are placed in the transparent acrylic box on the upper right side, and the edge computer with the lens module on the lower right side is connected to the control box with the PLC on the left side through Ethernet. First, the green light indicates the normal operation of the simulated production line, as shown in Figure 11(b). When the candle in the acrylic box is lit and the edge computer sends a fire extinguishing command to the PLC, the yellow light of the control box on behalf of the fire extinguishing action lights up. Finally, when the candle is extinguished, the edge computer informs that the fire alarm is over, and the red light on the control box that represents the stop of operation lights up. Figure 11(c) shows that the edge computer reports that the flame is detected and the control box enters the fire extinguishing action; and the Line Notify report of the fire extinguishing action is sent via the edge computer. Figure 11(d) shows that after the candle is extinguished, the edge computer informs the control box that the fire alarm is over, and the fire extinguishing action can be stopped. The control box reports that the extinguishing action is over after receiving the instruction and

enters the shutdown state; finally, it also reports that the extinguishing action has been completed through Line Notify. Finish.

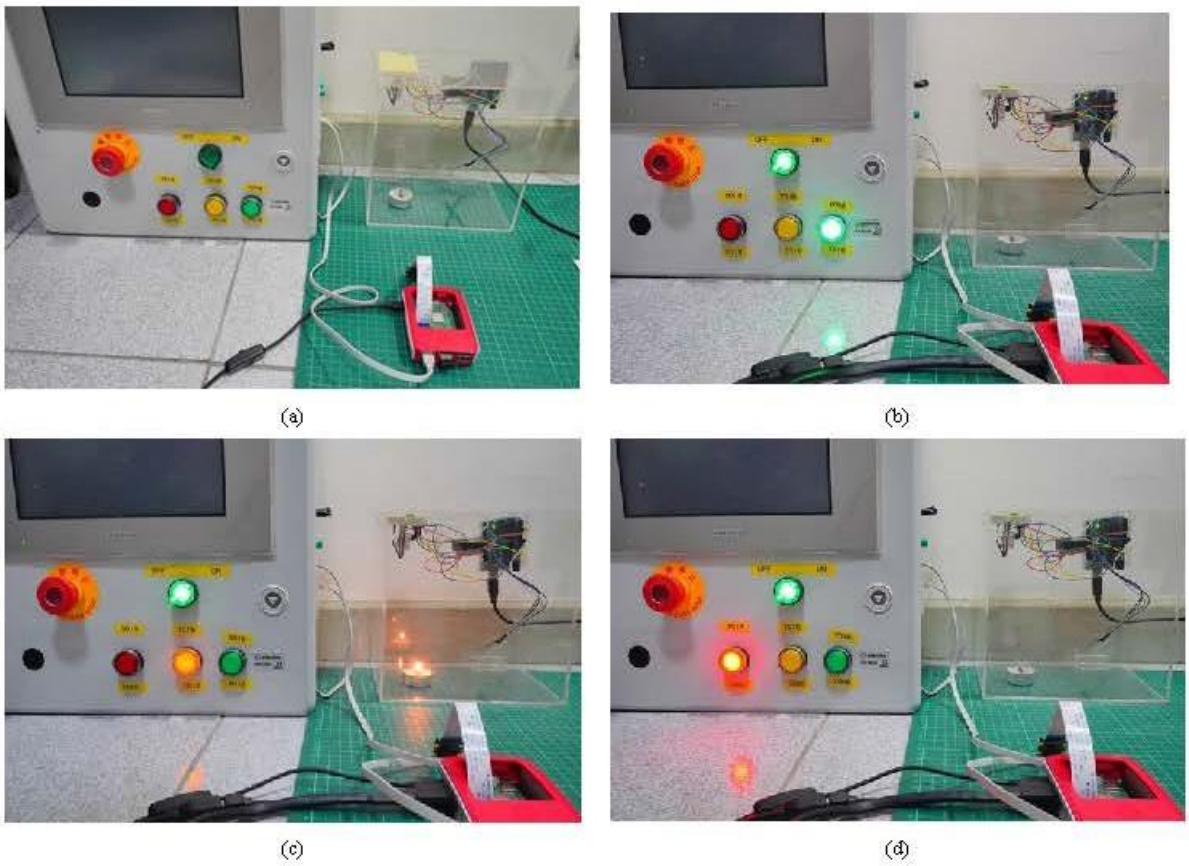

Fig. 11. (a) The edge computer controls the PLC fire extinguishing field (b) The control box operates normally (c) The control box starts fire extinguishing (d) The control box stops action

CONCLUSION

This research designed a set of intelligent automated industrial fire protection system, using a Bluetooth sensor designed in this research that is easy to install and free of settings. Through the Bluetooth broadcast mode, it continuously sends environmental data to the edge computer. The edge computer with smart judgment and smart control functions will collect data from Bluetooth sensors and perform algorithm calculations to determine whether the environment has drastically changed. Processing, even in the event of a fire, can autonomously control the PLC to extinguish and protect the production line equipment.

References

[1] A. Imteaj, T. Rahman, M. K. Hossain, M. S. Alam and S. A. Rahat, "An IoT based fire alarming and authentication system for workhouse using Raspberry Pi 3," *2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 2017, pp. 899-904, doi: 10.1109/ECACE.2017.7913031.

[2] S. Kabir, M. E. H. Riyad, U. T. Tanisha, M. H. R. Bhuiyan, T. R. Toha and S. M. M. Alam, "Automatic Fire Monitoring System in Textile Industry," *2020 11th International Conference on Electrical and Computer Engineering (ICECE)*, 2020, pp. 471-474, doi: 10.1109/ICECE51571.2020.9393112.

[3] M. Song et al., "Design of Distributed Factory Fire Alarm System," *2020 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS)*, 2020, pp. 132-135, doi: 10.1109/ICPICS50287.2020.9202068.

[4] C. Bhawesh and N. Gouthamkumar, "Advanced Integrated Fire Controlling System to Reduce Casualty and Minimize Deaths," *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, 2020, pp. 472-476, doi: 10.1109/ICACCCN51052.2020.9362770.

[5] Suresh S., Yuthika S. and G. A. Vardhini, "Home based Fire Monitoring and warning system," *2016 International Conference on ICT in Business Industry & Government (ICTBIG)*, 2016, pp. 1-6, doi: 10.1109/ICTBIG.2016.7892664.
[6] J. Ehlenbröker, U. Mönks, D. Wesemann and V. Lohweg, "Condition monitoring for hazardous material storage," *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*, 2014, pp. 1-4, doi: 10.1109/ETFA.2014.7005264.
[7] W. Chanthakhot and K. Ransikabum, "Numerical Simulation for Fire Emergency Planning in a Home Appliances Factory," *2019 Research, Invention, and Innovation Congress (RI2C)*, 2019, pp. 1-5, doi: 10.1109/RI2C48728.2019.8999928.
[8] M. S. A. Azmil, N. Ya'acob, K. N. Tahar and S. S. Sarnin, "Wireless fire detection monitoring system for fire and rescue application," *2015 IEEE 11th International Colloquium on Signal Processing & Its Applications (CSPA)*, 2015, pp. 84-89, doi: 10.1109/CSPA.2015.7225623.
[9] M. -H. Wang, S. -D. Lu, P. -Y. Ho, S. -E. Wei and C. -C. Hsieh, "Applying Internet of Things (IoT) Technology to Automatic Fire-extinguishing System in Machine Rooms," *2020 International Symposium on Computer, Consumer and Control (IS3C)*, 2020, pp. 17-18, doi: 10.1109/IS3C50286.2020.00011.
[10] S. Sharma, K. Chand, D. Sharma and P. Guha, "Development of an early detection system for fire using Wireless Sensor Networks and Arduino," *2018 International Conference on Sustainable Energy, Electronics, and Computing Systems (SEEMS)*, 2018, pp. 1-5, doi: 10.1109/SEEMS.2018.8687333.
[11] J. Razavi and N. Brennan, "A novel application of Bluetooth technology for detection of forest fires," *2016 IEEE Conference on Wireless Sensors (ICWiSE)*, 2016, pp. 66- 70, doi: 10.1109/ICWISE.2016.8188544.
[12] T. Islam, H. A. Rahman and M. A. Syrus, "Fire detection system with indoor localization using ZigBee based wireless sensor network," *2015 International Conference on Informatics, Electronics & Vision (ICIEV)*, 2015, pp. 1-6, doi: 10.1109/ICIEV.2015.7334000.
[13] A. R. Hutauruk, J. Pardede, P. Aritonang, R. F. Saragih and A. Sagala, "Implementation of Wireless Sensor Network as Fire Detector using Arduino Nano," *2019 International Conference of Computer Science and Information Technology (ICoSNIKOM)*, 2019, pp. 1-4, doi: 10.1109/ICoSNIKOM48755.2019.9111537.
[14] M. R. Habib et al., "Quick Fire Sensing Model and Extinguishing by Using an Arduino Based Fire Protection Device," *2019 5th International Conference on Advances in Electrical Engineering (ICAEE)*, 2019, pp. 435-439, doi: 10.1109/ICAEE48663.2019.8975538.
[15] M. S. Bin Bahrudin, R. A. Kassim and N. Buniyamin, "Development of Fire alarm system using Raspberry Pi and Arduino Uno," *2013 International Conference on Electrical, Electronics and System Engineering (ICEESE)*, 2013, pp. 43-48, doi: 10.1109/ICEESE.2013.6895040.
[16] R. I. Rashid, S. M. Rafid and A. Azad, "An Automated Fire Suppression Mechanism Controlled using an Arduino," *2018 IEEE Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER)*, 2018, pp. 49-54, doi: 10.1109/DISCOVER.2018.8674078.

Correlation between Overall Equipment Effectiveness (OEE) and Failure Mode and Effect Analysis (FMEA) at The Semi-Automatic Assembly Lines

Péter DOBRA[a]1, János JÓSVAI [b]

[a] *Doctoral School of Multidisciplinary Engineering Sciences, Széchenyi István University, Egyetem tér1, 9026 Győr, Hungary, e-mail: dobra.peter@sze.hu*
[b] *Department of Vehicle Manufacturing, Széchenyi István University, Egyetem tér1, 9026 Győr, Hungary, e-mail: josvai@sze.hu*

Abstract

In the industrial environment, continuous improvement of manufacturing and assembly processes is essential. Higher and faster customer's expectations have resulted in an increased complexity of manufacturing systems. Data collection has become easier and simpler, which has made it possible to develop new methods while conventional ones can be operated more efficiently. This paper deals with the relationship between two traditional methods such as Overall Equipment Effectiveness (OEE) and Failure Mode and Effect Analysis (FMEA). Firstly, a literature review demonstrates scientific relevance. Secondly, the correlation between the contributor of OEE (availability, performance, quality) and the components of FMEA (severity, occurrence and detection) is revealed. The complex analysis includes more than seven years real data from an automotive semi-automatic assembly line.

Keywords: OEE, FMEA, assembly line, correlation

1. INTRODUCTION

Manufacturing companies are under increasing pressure to raise productivity and performance levels. Enterprises systematically develop their manufacturing and assembly processes to meet customer expectations more cost-effectively. Higher and faster customer's expectations have resulted in an increased complexity of manufacturing systems. Data collection has become easier and simpler, which made the development of new methods possible. In addition, traditional methods can be operated more efficiently. Nowadays, we have larger set of raw data than we are able to process and evaluate parallel with ongoing operations.

Most of the automotive companies and suppliers use the Overall Equipment Effectiveness (OEE) indicator to track their effectiveness. This Key Performance Indicator (KPI) can be increased by various methods among others lean manufacturing, six sigma, line balancing and simulation. Different quality improvement tools such as Failure Mode and Effect Analysis (FMEA), cause and effect diagram, why-why analysis, failure tree analysis can also influence the current value of OEE. This article investigates the correlation between the contributor of OEE (availability, performance, quality) and the components of FMEA (severity, occurrence and detection).

* Corresponding author.

2. Literature review

In the scientific literature the concept of OEE and FMEA are extensive. According to Corrales et al. in terms of OEE more than 850 papers were published between 1996 and 2020 [1]. FMEA is one of the most effective fault analysis procedures in engineering field. Both methods are standardized based on the following:

- OEE (at the area of semiconductor manufacturing):
  - SEMI E10, Specification for Definition and Measurement of Equipment Reliability, Availability, and Maintainability (RAM)
  - SEMI E79, Specification for definition and measurement of equipment productivity
- FMEA:
  - BS EN IEC 60812:2018, Failure Modes and Effects Analysis (FMEA and FMECA)

2.1. The concept of Failure Mode and Effect Analysis

Within automotive industry, Advanced Product Quality Planning (APQP) is a standard method. This framework includes quality-oriented procedures and tools within the product creation process. One such tool is the FMEA, which is a highly structured preventive quality method. FMEA is based on former similar products and manufacturing processes [2]. FMEA is a systematic tool that identifies and evaluates potential failure modes and effects of entire production system in various phases of the product life cycle. The main purpose of the FMEA is to take actions to reduce or eliminate potential failures, starting with the highest-priority ones [3]. Another purpose is to provide a perfect product or service to the customer. The method is widely used in automotive industry, sector of aerospace, electronics, military, plastics, power plant, software, food industry, area of medical and pharmaceutical [4].

Based on Mil-P 1629 document FMEA method was introduced at the United States Military in 1949 [5]. It is used by the aerospace industry since the 1960s. In the 1970s Ford Motor Company introduced FMEA to the automotive industry. It was developed and standardized by the Auto Industry Action Group (AIAG) in the 1980s [6].

FMEA is characterized in the following ways:

- qualitative technique of continuous improvement [7]
- proactive failure analysis quality tool [8]
- analytical quality planning tool dedicated potential failure modes [3]
- structured approach to improve quality, reliability and safety [9]

There are many types of FMEA among others: Concept FMEA (CFMEA), Design FMEA (DFMEA), Process FMEA (PFMEA), Maintenance FMEA (MFMEA), Logistics FMEA (LFMEA). Process FMEA has two part such as Manufacturing FMEA and Assembly FMEA. The main advantages of using the FMEAs are reduction of warranty costs, improvement of the quality and reliability of the products, increasing the safety in the operation, reduction of the time needed from the project phase to the market launch [10].

FMEA determines the risk priority number (RPN), which is given by the product of the severity (S), occurrence (O) and detection (D) of the failure.

$$RPN = S \times O \times D \qquad (1)$$

The value of the three factors is estimated to vary between 1 and 10 so the RPN will be between 1 and 1000. In case of higher RPNs the operative team must undertake efforts to reduce this calculated risk through corrective actions.

Heutmann and Schmitt used the risk priority number for disturbance prediction at the synchronized individual production [11]. According to Fabarisov et al. FMEA, Failure Mode Effects and Criticality Analysis (FMECA) and Hazard and Operability Analysis (HAZOP) are top-level qualitative methods that require numerical input from quantitative methods [12]. Subriadi and Najwa presented an improved FMEA model based on the traditional FMEA. The developed model was validated at the field of information technology risk assessment [13]. Sutrisno et al. developed a modified FMEA as means to access the criticality of waste in maintenance operations. They defined a metric which was called Waste Priority Number (WPN) and used as surrogate of maintenance waste risk. WPN index is having the similar function like the RPN in conventional FMEA [14].

The weaknesses of FMEAs are as follows: difficult to find the potential root causes, non-linear 1–10 priority scales,
evaluating risk factors accurately, subjectivity human error, time-consuming, the importance level of parameters which are similar, main focus on the RPN. Factors are not generated by measurement but by determination, so their values are not very reliable [6, 13]. Sankar and Prabhu found that the different RPN values are not evenly distributed on a scale of 1 to 1000. At the beginning of the scale, the RPN number can take values in many more combinations [15]. Bowles and Peláez first applied the fuzzy method to the FMEA procedure to overcome the disadvantages of FMEA [16]. According to Ishak et al. Fuzzy FMEA is an efficient method in determining priority numbers of risk of failure of products to be identified prior to repair actions [4].

2.2. Overall Equipment Effectiveness (OEE) as a Key Performance Indicator (KPI)

According to the scientific literature several Key Performance Indicator (KPI) are used to measure production efficiency. In the production domain, Overall Equipment Effectiveness (OEE) is a widely used and most famous performance indicator [17]. OEE was introduced by Nakajima when the Total Productive Maintenance (TPM) concept was presented in 1988. It is the product of the availability, performance and quality of the machine or assembly line. Under ideal circumstances availability is higher than 90%, performance is higher than 95% and quality rate is higher than 99%. According to this conditions OEE percentage is higher than 84.6% [18]. Hansen has determined excellent OEE percentage values for different production types, in case of batch production it has the value of 85% [19].

The basic formula for calculating OEE is written as:

$$OEE = A \times P \times Q \ [\%] \qquad (2)$$

where: A - availability [%]; P - performance [%]; Q - quality [%].

The value of OEE can be increased in numerous ways among others:

- Lean manufacturing tools
  - 5S, Andon, Heijunka, Kaizen, Kanban, Line Balancing, One-Piece-Flow, Poka Yoke, Pull System, Shop Floor Management, Single Minute Exchange of Die (SMED), Standard Operating Procedure (SOP), Total Productive Maintenance (TPM), Value Stream Mapping (VSM), Visual Management, Yamazumi
- Quality tools
  - 5 Why, 5W2H, Cause and Effect Diagram, Check Sheet, Control Chart, Histogram, Pareto Chart, Plan Do Check Act (PDCA), Quick Response Quality Control (QRQC), Root Cause Problem Solving (RCPS), Statistical Process Control (SPC)
- Other methods

– Best Practices, Bottleneck Detection, Continuous Improvement (CI), Data Mining, Enterprise Resource Planning (ERP), Failure Mode and Effect Analysis (FMEA), Good Idea System, Manufacturing Execution System (MES), Methods Time Measurement (MTM), Performance Management, Simulation, Six Sigma, Teamwork, Waste Hunting [20]

According to Ales et al. the detection of critical points in production lines and taking measures to increase the reliability and efficiency is what OEE is calculated for [21].

2.3. FMEA and OEE connection

Ahire and Relkar examined a number of hypotheses regarding the relationship between OEE and FMEA [22]. Although these studies are based on data from a very short period of time. For higher OEE percentage and lower RPN value the companies use the principle of zero-defect manufacturing (ZDM). The main approaches are:

- Monitoring of process parameters and in-line measurement
- Collaborative manufacturing
- Continuous quality control and focus on prevention
- On-line predictive maintenance
- Data storage and analytics
- Re-configuration and re-organization of production
- Re-scheduling of production [23, 24].

Garza-Reyes et al. analyzed the relationship between OEE and process capability [25]. Relkar and Nandurkan attempted to predict the OEE by using Design of Experience (DOE). Their research work indicates that OEE will be significantly improved if focus is given on performance rate improvement [26]. Okpala et al. made a descriptive analysis and used Pearson correlation for predicting OEE at a pharmaceutical company. The analysis shows that performance and quality are significant in predicting the Overall Equipment Effectiveness, unlike Availability that is not significant [27].

3. Correlation between OEE and FMEA

Although OEE and FMEA has been known for decades their relationship based on several years of data has not been described so far. In this chapter the focus is on the Process FMEA (PFMEA). In the industrial practice, the PFMEA will never be finished unless the product is removed from production. This will result in several versions being made of the FMEA. When a change is made to the design, application, environment, material, or any process step of a product, the FMEA must be updated (e. g. drawing modification, measurement method modification, etc.). In the case of customer complaints, the manufacturing, assembly or inspection processes will change, so the PFMEA will also be modified. Over a longer period of production, there are plenty of changes that affect OEE and FMEA. New product variants are launched, change in the engineering and operator staff, increase of the production and assembly experiences, etc.

3.1. Detailed analysis at the semi-automatic assembly line using Pearson correlation

Pearson correlation is used to identify potential dependencies [28]. The main question is whether there is a relationship between two or more quantitative variables. The general formula used is as follows:

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} \quad (3)$$

The possible correlations between OEE and FMEA is analyzed at an automotive seat structure semi-automatic assembly line (Fig. 1.). The line has been producing since 2014 and assembling metal seat structure for automotive industry. Assembly line operates in two shifts with twelve operators and with eight automatic stations. The blue arrow shows the direction of material flow through twenty workplaces. Since the beginning of production this line has been producing the same product, there is no significant difference between the product variants. Although a total of one hundred types are assembled in small batch sizes. The average batch number is between 300 and 500 units.

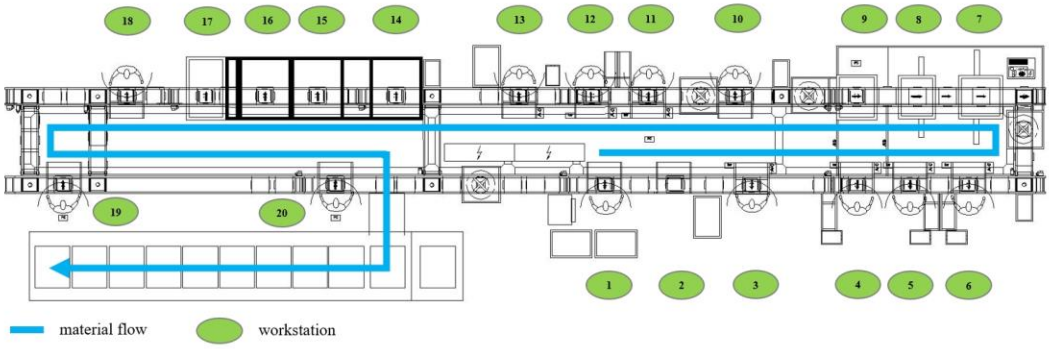

Fig. 1. Seat structure semi-automatic assembly line

Based on Manufacturing Execution System (MES), every OEE data is available for different periods. The exact availability, performance and quality values can be determined. Data from 87 months (more than 7 years) were used in the case study. In terms of FMEA, 40 versions have been created within 7 years.

The correlation analysis method is as follows:

- Determination of OEE components with the support of MES on monthly basis
- Calculating monthly OEE percentage
- Overview of all FMEA versions
- Determining the mean values for severity, occurrence and detection for each version
- Determining the average RPN values for each version
- Comparison of OEE contributors and RPN components on monthly basis (Fig. 2.)
- Correlation analysis, results
- Conclusions

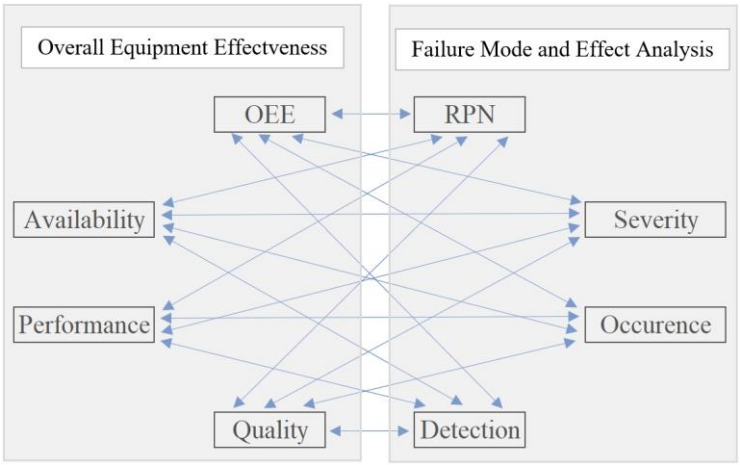

Fig. 2. Comparison of OEE contributors and FMEA components

To interpret the effect size *r* of each factor combination Cohen provides the limits shown in Table 1. [29].

Table 1. Correlation classes

Correlation strength	Effect size
No correlation	$r < 0.1$
Small correlation	$0.1 \leq r < 0.3$
Medium correlation	$0.3 \leq r < 0.5$
Strong correlation	$0.5 \leq r$

Applying the method mentioned above, Table 2. shows the correlations between OEE and RPN. All possible relationships and value effects (r) are described. Calculation was supported by Excel.

Table 2. Correlation between OEE and RPN elements

Relationship	Value of effect (r)
Performance - Risk Priority Number	-0.854231
Performance - Detection	-0.802474
Overall Equipment Effectiveness - Risk Priority Number	-0.785982
Quality - Risk Priority Number	-0.764704
Performance - Occurrence	-0.752277
Overall Equipment Effectiveness - Detection	-0.704877
Quality – Detection	-0.695907
Overall Equipment Effectiveness – Occurrence	-0.646216
Availability - Risk Priority Number	-0.639994
Quality – Occurrence	-0.635789
Availability - Detection	-0.577365
Availability – Occurrence	-0.500222
Performance – Severity	0.550503
Availability - Severity	0.554767
Quality – Severity	0.570743
Overall Equipment Effectiveness - Severity	0.641808

In order to be able to analyze the strength of the relationship between the quantitative variables, the effect values had to be sorted. Fig. 3. shows the ranking and the value of the effect.

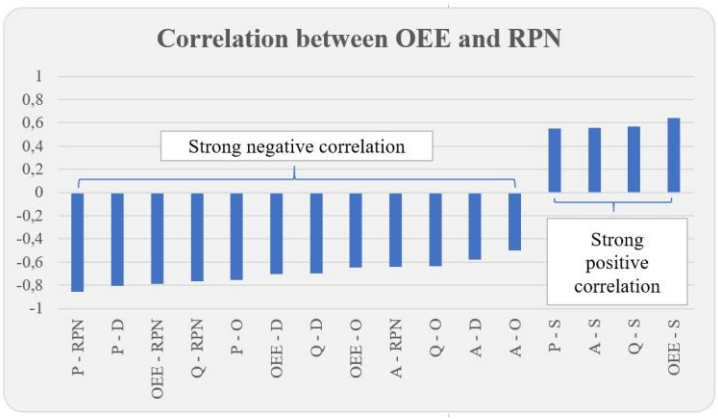

Fig. 3. Ranking and correlation between OEE and RPN

The following were established after the analysis:

- Based on Cohen's suggested classes, there is either a strong negative or strong positive correlation between all factors
- TOP 3 strong negative correlation:
  - as performance increases, the RPN number decreases
  - as performance increases, detection average value decreases
  - as OEE percentage increases, the RPN number decreases.
- TOP 3 strong positive correlation:
  - as OEE percentage increases, the severity average value also increases
  - as quality percentage increases, the severity average value also increases
  - as availability percentage increases, the severity average value also increases.
- The RPN number continuously decreases during the analyzed period. At first faster (ramp-up period) later on slower.
- OEE percentage continuously increasing during production ramp-up period, later increasing slower and shows stabilized period. Fig. 4. shows the monthly OEE and average RPN values from Q2/2014 until Q2/2021.

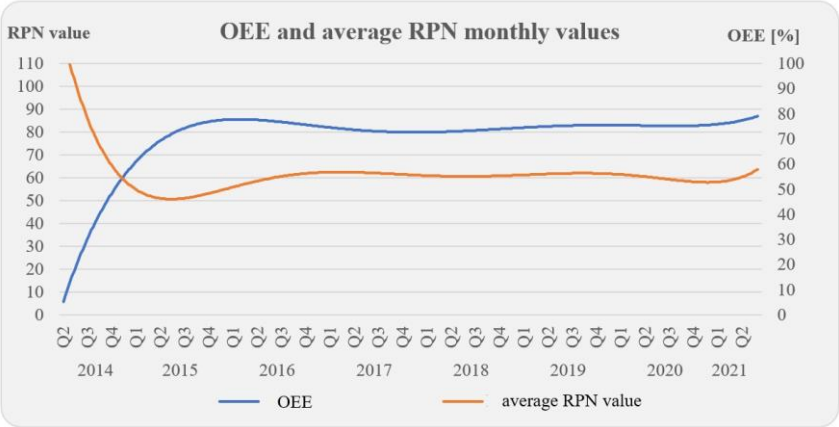

Fig. 4. Ranking and correlation between OEE and RPN

Based on the analysis of long-term data, the increase in the percentage of OEE is closely related to the decrease in the monthly average value of RPN. In the case of semi-automatic assembly lines, the systematic use of FMEA results in improved efficiency.

CONCLUSIONS

Failure Mode and Effect Analysis (FMEA) and Overall Equipment Effectiveness (OEE) are widespread methods in the automotive industry. The possible correlations between OEE and FMEA is analyzed at an automotive seat structure semi-automatic assembly line. Data from 87 months (more than 7 years) were used in the case study. In terms of FMEA, 40 versions have been created within 7 years. The correlation between the contributor of OEE (availability, performance, quality) and the components of FMEA (severity, occurrence and detection) is revealed by using Pearson correlation. TOP 3 strong negative and positive correlation were identified. Based on the analysis of long-term data, the increase in the percentage of OEE is closely related to the decrease in the monthly average value of RPN.

References

[1] Corrales, L. C., Lambán, M. P., Korner, M. E. H., Royo, J. (2020). Overall Equipment Effectiveness: Systematic literature review and overview of different approaches. *Applied Science, 10, 6469.*

[2] Kiefer, J., Allegretti, S., Breckle, T. (2017). Quality- and lifecycle-oriented production engineering in automotive industry. *Procedia CIRP, 62, 446-451.*

[3] Waghmare, S. N., Raut, D. N., Mahajan, S. K., Bhamare, S. S. (2014). Failure Mode Effect Analysis and Total Productive Maintenance: A review. *International Journal of Innovate Research in Advanced Engineering, 1 (6), 183-203.*

[4] Ishak, A., Siregar, K., Ginting, R., Mainik, A. (2020). Implementation Statistical Quality Control (SQC) and Fuzzy Failure Mode and Effect Analysis (FMEA): A systematic Review. *IOP Conference Series: Materials Science and Engineering.*

[5] Procedures for Performing a Failure Mode, Effect and Criticality Analysis. MIL-P-1629: U.S. Department of Defense (1949).

[6] Juhász, L. (2016). A hibamód és hatáselemzés irodalmi áttekintése (in Hungarian). *Repüléstudományi Közlemények, 3, 95-108.*

[7] Dias, J., Nunes, E., Sousa, S. (2020). Productivity improvement of transmission electron microscopes – A case study. *Procedia Manufacturing, 51, 1559-1566.*

[8] Chong, K. E., Ng, K. C., Goh, G. G. G. (2015). Improving Overall Equipment Effectiveness (OEE) through integration of Maintenance Failure Mode and Effect Analysis (Maintenance-FMEA) in a semiconductor manufacturer: A case study. *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Singapore, 1427-1431.*

[9] Deák, Cs. (2005). *Minőségmenedzsment* (in Hungarian). Miskolc: Bíbor kiadó.

[10] Sharma, K. D., Strivastava, S. (2018). Failure Mode and Effect Analysis (FMEA) implementation: A literature review. *Journal of Advance Research in Aeronautics and Space Science, 5, (1-2) 1-17.*

[11] Heutmann, T., Schmitt, R. H. (2019). Predictive control of a synchronized individual production. *Procedia CIRP, 81, 69-74.*

[12] Fabarisov, T., Yusupova, N., Ding, K., Morozov, A., Jauschek, K. (2020). Model-based stochastic error propagation analysis for Cyber-Physical Systems. *Acta Polytechnica Hungarica, 17 (8), 15-28.*

[13] Subriadi, A. P., Najwa, N. F. (2020). The consistency of Failure Mode and Effect Analysis (FMEA) in information technology risk assessment. *Heliyon, 6 (1), e03161.*

[14] Susrisno, A., Gunawan, I., Tangkuman, S. (2015). Modified Failure and Effect Analysis (FMEA) Model for accessing the risk of maintenance waste, *Procedia Manufacturing, 4, 23-29.*
[15] Sankar, R. N., Prabhu, B. S. (2001). Modified approach for prioritization of failures in a system failure mode and effect analysis. *International Journal of Quality and Reliability Management, 18 (3), 324-336.*
[16] Bowles, J. B., Peláez, C. E. (1995). Fuzzy logic prioritization of failures in a system failure mode, effect and criticality analysis. *Reliability engineering and system safety, 50 (2), 203-213.*
[17] Kurve, M., Harlin, U., Hallin, M., Söderlund, C., Berglund, M., Florin, U. (2019) Designing visual management in manufacturing from user perspective. *Procedia CIRP, 84, 886-891.*
[18] Nakajima, S. (1988). *Introduction to TPM: Total Productive Maintenance.* Cambridge: Productivity Press.
[19] Hansen, R. C. (2001). *Overall Equipment Effectiveness: A powerful production / maintenance tool for increased profits.* New York: Industrial Press.
[20] Dobra, P. (2021). Towards 100% Overall Equipment Effectiveness (OEE) at semi-automatic assembly lines. *28th International Conference on System Engineering (paper in production).* Wroclaw, Poland.
[21] Ales, Z., Pavlu, J., Legét, V., Mosna, F., Jurca, V. (2019). Methodology of Overall Equipment Effectiveness calculation in the context of Industry 4.0 Environment. *Ekspolatacja I Niezawodnose – Maintenance and Reliability, 21 (3), 411-418.*
[22] Ahire, C. P., Relkar, A. S. (2012). Correlating Failure Mode and Effect Analysis (FMEA) and Overall Equipment Effectiveness (OEE). *Procedia Engineering, 38, 3482-3486.*
[23] Lindström, J., Lejon, E. Kyösti, P., Mecella, M., Heutelbeck, D., Hemmje, M., Sjödahl, M., Birk, W, Gunnarsson, B. (2019) Towards intelligent and sustainable production systems with a Zer-Defect Manufacturing approach in an Indiustry 4.0 context. Procedia *CIRP, 81, 880-885.*
[24] Eger, F., Coupek, D., Caputo, D., Colledani, M., Penalva, M., Ortiz, J. A., Freiberger, H., Kollegger, G. (2018). Zero defect manufacturing strategies for reduction of scrap and inspection effort in multi-stage production system. *Procedia CIRP, 67, 369-373.*
[25] Garza-Reyes, J. A., Eldridge, S., Barber, K. D., Soriano-Meyer, H. (2010). Overall equipment effectiveness (OEE) and process capability (PC) measures. *International Journal of Quality and Reliability Management, 27 (1), 48-62.*
[26] Relkar, A. S., Nandurkar, K. N. (2012). Optimizing and analyzing Overall Equipment Effectiveness (OEE) through Design of Experiments (DOE). *Procedia Engineering, 38, 2973-2981.*
[27] Okpala, C. C., Anozie, S. C., Mgbemena, C. E. (2020). The optimization of overall equipment effectiveness factors in a pharmaceutical company. *Heliyon, 6, e03796.*
[28] Denkena, B., Dittrich, M. A., Wilmsmeier, S. (2019). Automated production data feedback for adaptive work planning and production control. *Procedia Manufacturing, 28, 18-23.*
[29] Cohen, J. (1988). *Statistical power analysis for behavioral sciences.* Hillsdale NJ, Lawrence Erlbaum Associates.

A Data-driven Solution for The Dynamic Capacitated Arc Routing Problem

Zsuzsanna NAGY[a1], Ágnes WERNER-STARK [b], Tibor DULAI [c]

University of Pannonia, Faculty of Information Technology, Department of Electrical Engineering and Information Systems, Egyetem u. 10, 8200 Veszprém, Hungary
[a] nagy.zsuzsanna@virt.uni-pannon.hu, [b] werner.agnes@virt.uni-pannon.hu, [c] dulai.tibor@virt.uni-pannon.hu

Abstract

The Capacitated Arc Routing Problem (CARP) is a combinatorial optimization problem that requires determining the least cost route plans on a graph for vehicles subject to some constraints. Dynamic CARP (DCARP) is a variant of CARP that considers dynamic changes in the problem during the execution of the plan. Most of the existing works only consider one type of dynamic event and when such event occurs, all the route plans are re-planned. Nowadays, technologies can provide nearly instant updates on information about the roads (e.g., travel time and road closure). Furthermore, data can be gathered in real time about the traversed and served road sections. In this work, a data-driven DCARP framework is proposed, that can handle all the events that can make the original schedule infeasible or obsolete. Furthermore, a novel rerouting algorithm (RR1) is also introduced, that aims to find the least cost route plans by utilizing a travel and service log and rerouting at most one route plan. The results showed that RR1 can give solutions fast and the given solutions in around half of the cases have the same cost as the solutions given by the currently most accurate routing algorithm.

Keywords: Dynamic capacitated arc routing problem, Event stream, Virtual task strategy

1. INTRODUCTION

The Capacitated Arc Routing Problem (CARP) is a combinatorial optimization problem that requires determining the least cost route plan on a road map for vehicles subject to some constraints. The road map, which is represented as a graph, contains a set of vertices that are connected by edges and arcs (i.e., directed edges). Each edge and arc have a travel cost, and some of them also have demands required to be served by vehicles and service cost, which is the cost of serving the demands (that includes the traveling cost too). The edges and arcs with demands are called edge tasks and arc tasks or simply tasks. By optimizing CARP, a number of vehicles with limited capacities are assigned to the graph to serve all tasks spending the lowest possible total cost.

CARP has many applications in the real world, such as urban solid waste collection [1] [2] and winter gritting (salting the streets) [3][4], so it always has been a popular research topic. Since CARP is NP-hard [5], exact methods are only applicable to small-size instances. For this reason, heuristics and metaheuristics are considered most of the times in the literature [6][10].

CARP belongs to the family of arc routing problems (ARPs) which has many variants. ARP itself is a variant of vehicle routing problems (VRPs) [14]. Interested readers can read about the existing variants of ARPs in the annotated bibliographies [11] [12] that collected and cataloged them according to their specifics up to 2017, or in the most recent literature review [13] that was published in 2021.

[1] Corresponding author.

Dynamic CARP (DCARP) is a variant of CARP that considers dynamic changes in the problem during the execution of the plan. In real-life applications, dynamic changes may happen when vehicles are in service, thus influencing the execution of the route plan. If a dynamic event drastically affects the current schedule (e.g., a new task appears or a vehicle breaks down before it served all the tasks on its route plan), rescheduling of the route plans is necessary. Despite its importance, the number of studies that consider dynamic changes during the vehicles' service is relatively small [15] - [23], and the number of studies that consider more than two events is even smaller [16], [23]. For a more detailed comparison of these works, see Table 3. in Section 4.3.

Recently a framework was introduced, which makes it possible to use static CARP algorithms to reschedule the route plans of a DCARP instance. Using this framework was proven to be more effective than using the existing DCARP solutions, which are usually limited to handle only a few kinds of dynamic events and also suffer from slow runtime [23]. However, all these solutions allow only complete rerouting of the current solution, which can be time consuming and can also change the route plans of several vehicles, which may confuse the drivers. There is a work on VRP that for instance, in case of vehicle breakdown event, modifies the route plan of only one vehicle (to serve the unexecuted tasks of the broken down vehicle), and was able to achieve good results [24].

The aim of this work is to create a solution for DCARP that

- utilizes the available data on the vehicles' activity (i.e., it is data-driven), since there can be differences between the route plans and the execution of these plans;
- can handle the critical events, the events that can make the original schedule infeasible and/or obsolete and can make rerouting necessary;
- reroutes only when it is necessary and during the rerouting process tries to find the least cost route plans but it reroutes at most as many route plans of the current solution as it is necessary to make the solution feasible again.

In this work, the possible events that change the instance were collected, then analyzed. Based on the results, three event handling algorithms were developed that use the virtual task strategy [23] and a novel rerouting algorithm (RR1) that reroutes at most one route plan and (depending on the case) modifies at most two route plans of the current solution, while aiming to find the least cost route plans. If a new task appears, then that one task has to be added to the solution. Therefore, in such a case, modifying one route plan is sufficient to make the solution current again. If the demand of a task increases in an extent that the total demand of the tasks in the route plan exceeds the capacity of the serving vehicle, or a vehicle with at least one unexecuted task breaks down, then at least one unexecuted task has to be moved into another plan. Therefore, in such cases at least two route plans have to be modified to make the solution feasible and current again. However, only one route plan is rerouted (the one that takes over the task(s) in question). The algorithms use the travel and service log of the vehicles. The implemented algorithms were tested on a wide range of DCARP instances and were compared with existing solutions.

The paper is constructed as follows. In Section 2, the problem formulation of DCARP for this work is introduced. In Section 3, the data sources for DCARP are described. In Section 4, the possible dynamic events that change the instance are analyzed. In Section 5, the developed algorithms (the event handling algorithms and the rerouting algorithm) and the modified virtual task constructor algorithm are introduced. In Section 6, the methods are described, that were used to generate events for testing the algorithms. In Section 7, the results of the experiments are presented. Finally, in Section 8, the paper is concluded.

2. PROBLEM FORMULATION OF DCARP

In this section, the problem formulation of DCARP for this work is introduced. It is mainly based on the mathematical formulation introduced for DCARP in [23], but there are some differences.

2.1. Scenario

A DCARP scenario is composed of a series of DCARP instances $I = \{I_0, I_1, \ldots, I_m, \ldots, I_M\}$, where M is the number of dynamic events that occurred and changed the current DCARP instance (i.e., created a new DCARP instance). Each DCARP instance corresponds to a problem state, which contains all the information regarding the state of the road map and vehicles involved in the routing problem, and highly depends on the previous instance and the execution of the solution.

The initial problem instance I_0 is fundamentally a static CARP, since in it all vehicles are located at the depot. The initial solution can be obtained from a static CARP solver, then this solution can be started to be executed. During the execution, due to occurrence of dynamic events some changes happen at random time points, thus changing the problem instance and requiring a new solution to be found. Vehicles then continue to serve tasks from the positions they had stopped (stop points) following the new solution. DCARP terminates when all tasks are performed, and all (working) vehicles have returned to the depot.

In a DCARP scenario, usually, the main objective is to achieve a schedule plan for each DCARP instance, which has an execution cost as low as possible. In this work, the main objective is to achieve a schedule plan that keeps the number of modified route plans low (at most one) between the consecutive DCARP instances, and the secondary objective is to keep the total cost as low as possible.

2.2. Instance

In the existing works on CARP, there is no consensus on the type of the graph; some works assume an undirected graph [9], others a directed graph [10], and another ones a mixed graph [6], [7]. To be able to handle all three types of graphs, in this work a mixed graph is assumed.

A static CARP instance I is usually defined the following way:

$$I = (V, E, A, E_R, A_R, v_0, Q, N, head, tail, dc, inv, id, dem, sc)$$

where

- (V, E, A) is the mixed graph with a set of vertices V, a set of (undirected) edges E, and a set of arcs (directed edges) A.
- $E_R \subseteq E$ is the set of required edges, $A_R \subseteq A$ is the set of required arcs and they jointly define the tasks. Although an edge task has two arcs, it should be served only once, in either direction.
- The tasks need to be served by a fleet of N identical vehicles with capacity Q, whose route starts and ends in the depot vertex v_0 ($v_0 \in V$).
- Each arc $a \in (E \cup A)$ has a head vertex $head(a)$, a tail vertex $tail(a)$ and a dead-heading or traversing cost $dc(a)$ (i.e., the cost of crossing the arc without serving it).
- Each arc task $t \in (E_R \cup A_R)$ in addition has a unique identifier $id(t)$, positive demand $dem(t)$ (it indicates the load necessary to serve the task) and service cost $sc(t)$ (i.e., the cost of crossing and serving the task, the traversing cost is included in this cost). Each arc task $t \in E_R$ (i.e., originally edge task) also has an inverse task $t_{inv} = inv(t)$, that is on the inverse arc (i.e., $head(t) = tail(t_{inv})$ and $tail(t) = head(t_{inv})$), but has the same demand and service cost (i.e., $dem(t) = dem(t_{inv})$ and $sc(t) = sc(t_{inv})$).

For a DCARP instance additional information is needed to be stored as well, such as the current location of the outside vehicles and their remaining capacities, which could be different for each vehicle. This problem can be fixed by the virtual task strategy introduced in [23], which converts a DCARP instance to a virtual static CARP instance by using ”virtual tasks”. A virtual task is an arc that connects the depot with the current location of the vehicle (i.e., its head vertex is v_0 and its tail vertex is a $v \in V$, assuming that every vehicle is currently located at a vertex at the time when an unexpected event occurs). Since it is “virtual”, it cannot be traversed, so it has an infinite traversing cost. Furthermore, since it is a “task”, it has demand and service cost. In [23], the service cost of a virtual task is defined as the minimal traversing cost from the depot to the current position of the vehicle, and the demand is defined as the difference between the maximum capacity and the remaining capacity of the vehicle. In this work, the service cost of the virtual task is defined as the total cost (traversing and serving cost) produced by the vehicle so far according to the travel and service log (Section 3.1), and the demand is defined as the total demand served by the vehicle so far also according to the log. Although the demand is calculated in a different way, the resulting value is the same (served demand + remaining capacity = maximum capacity, so served demand = maximum capacity - remaining capacity, where the remaining capacity refers to the capacity of the vehicle that still can be used to serve tasks, not taking into account the total demand of the unexecuted part of the current route plan). However, the service cost is very different.

In this work, a DCARP instance I_m is defined the following way:

$$I_m = (V, v_0, A, T, T_v, Q, H, H_f, R, R_e, rt, rv, head, tail, dc, inv, dem, sc, mdc)$$

where

- V is the set of vertices and v_0 ($v_0 \in V$) is the depot vertex.
- A is the set of arcs that contains each edge as two arcs (i.e., one arc for each direction). The arcs and the edges are easily distinguishable with the *inv* function, since it can only be interpreted for edges. For instance, if $e = (v_i, v_j)$ is an edge, then both e and $inv(e) = (v_j, v_i)$ are members of A, if $v_i, v_j \in V$.
- T is the set of tasks. Just like in A, the edge tasks are defined with two arc tasks. As it was mentioned before, for an edge task, only one of its arcs has to be served. The total number of tasks that need to be served is denoted by N_t.
- T_v is the set of virtual tasks. Since virtual tasks need to be updated for the outside vehicles that are still in service wherever a new instance is created, virtual tasks are preferred to be stored separately from normal tasks.
- H is the set of identifiers of the vehicles, which are identical with capacity Q.
- $H_f \subseteq H$ is a set of identifiers of the currently free vehicles, the vehicles that (based on the travel and service log) are currently in the depot. It is used during the event processing and route plan rerouting process to assign vehicles to new route plan(s).
- R is the set of identifiers of the route plans of the solution of the previous DCARP instance I_{m-1}.
- $R_e \subseteq R$ is a set of identifiers of the route plans where (permanently or temporarily) more tasks cannot be added, since the route is already finished (i.e., the vehicle which was executing it returned to the depot) or the vehicle which was executing it broke down. It is used during the rerouting process to prevent the used algorithm from adding tasks to these route plans. The routes that cannot serve tasks are kept in the solution for two reasons: (1) for the total cost computation and (2) if the route plan was executed by a vehicle, which broke down before serving all its tasks, and later it restarts, then it can continue serving its remaining tasks or other tasks before returning to the depot. Furthermore, (3) if the broken down vehicle have to be visited by another vehicle to take the goods, but this is not considered in this work, since the goods are homogeneous.
- $rt : R \rightarrow T_v$ is a function that assigns to the identifiers of the route plans their current virtual task. For instance, if x is a route plan identifier (i.e., $x \in R$), then $rt(x)$ is the current virtual task of x (i.e., $rt(x) \in T_v$).
- $rv : R \rightarrow H$ is a function that assigns to the identifiers of the route plans the identifiers of the vehicles that are executing them. One vehicle might be assigned to multiple route plans. In such case, the route plan with the greatest identifier value is the currently executed plan and the rest are the previously executed, already finished plans. Therefore, the identifier of the currently followed route plan of the vehicle with identifier $y \in H$ is $x_{curr} = \max\{ x \mid x \in R, rv(x) = y \}$, where x is an identifier of an existing route plan (that is element of the domain of the function rv). One route plan might not have a vehicle assigned to it yet, if currently there are no free vehicles available to execute the plan. In such cases $rv(x) = -1$, where x is an identifier of an existing route plan (i.e., $x \in R$).
- The functions *head*, *tail*, *dc*, *inv*, *dem*, and *sc* are defined the same way as in a static CARP instance.
- The function $mdc(v_i,v_j)$ returns the minimal total traversing cost from node v_i to node v_j, if $v_i, v_j \in V$. These values are calculated with Dijkstra's algorithm.

The optimization of DCARP aims to reschedule the plan to serve the tasks that were not served yet with minimal cost considering both outside and depot vehicles.

2.3. Solution

A solution for a DCARP instance can be expressed as a set of routes, where the routes are sequences of the tasks that need to be visited in the given order. The consecutive tasks are connected by the shortest paths (i.e., paths with minimal total dead-heading cost).

A solution for DCARP can be expressed as $S = \{r_1, r_2, \ldots, r_K\}$, where K is the number of routes. The task sequence represented by a given route r_k can be expressed as $r_k = (v_0, t_{k,1}, t_{k,2}, \ldots, t_{k,lk}, v_0)$, where the vehicle starts from and returns to the depot v_0, and l_k denotes the number of tasks served by route r_k. For outside vehicles, $t_{k,1}$ is a virtual task (i.e., $t_{k,1} \in T_v$).

2.4. Objective and constraints

The objective of DCARP is to minimize the total cost of the solution subject to the defined constraints. The total cost of a solution *S* is calculated with the following formula below (*1*), where (2) is the total dead-heading cost and (*3*) is the total service cost of route r_k.

$$TC(S) = \sum_{k=1}^{K} DC(r_k) + SC(r_k) \quad (1)$$

$$DC(r_k) = mdc\left(v_0, head\left(t_{k,1}\right)\right) + \sum_{i=1}^{l_k-1} mdc\left(tail\left(t_{k,i}\right), head\left(t_{k,i+1}\right)\right) + mdc\left(tail\left(t_{k,l_k}\right), v_0\right) \quad (2)$$

$$SC(r_k) = \sum_{i=1}^{l_k} sc(t_{k,i}) \quad (3)$$

A DCARP solution has to satisfy three constraints, which are the same constraints as in static CARP. Therefore, a solution *S* has to satisfy the following constraints:

- Each route served by one vehicle must end at the depot.
- Each task is served exactly once. Therefore, the total number of tasks served on each route has to be equal to the total number of tasks that have not been served yet N_t and the total number of virtual tasks $|T_v|$:

$$\sum_{k=1}^{K} l_k = N_t + |T_v|$$

  And a task cannot be served more than once, neither in the same route nor in another route:

$$t_{k,i} \neq t_{y,j}, \forall (k,i) \neq (y,j)$$

  where r_k and r_y are route plans within the same solution, $t_{k,i}$ is the *i*-th task in the route plan r_k, and $t_{y,j}$ is the *j*-th task in the route plan r_y.

- The total demand for each route served by one vehicle cannot exceed the vehicle's capacity. Therefore, the total demand for each route has to be smaller or equal to the capacity of the vehicle.

$$\sum_{i=1}^{l_k} dm(t_{k,i}) \leq Q, \qquad \forall k \in \{1, 2, \dots, K\}$$

3. EVENT STREAMS AND LOGS

In terms of process mining, an event log represents the observed behavior of a process, which contains a finite number of events. Every event has an activity and a case identifier (that identifies the context in which the activity was executed) recorded, but they can contain additional information as well (e.g., resource, which is the person/machine who executed the event). An event stream is defined as an (infinite) sequence of events, where an event (also called observable unit) is an n-tuple of a case identifier, an activity identifier and some additional information about the event (i.e., attributes of the event). From the observed events on the event stream, an event log can be constructed.

3.1. Travel and service event stream and log

In case of a (D)CARP instance, executing the route plans can be seen as a process where events are generated by vehicles when they start or finish traversing or serving a road segment. Therefore, the activity is traversing/serving an arc/edge of the graph and the resource is the vehicle. To make the processing of an event easier, the arc (for edges, the corresponding arc of the edge) and whether it was served or only traversed are stored in separate attributes. Therefore, an event (or an observable unit) is constructed as follows:

1. the time point when the traversing/serving of the road segment started/completed, it can be represented multiple ways;
2. the unique identifier of the vehicle, preferably an integer number;

3. the arc $a \in A$ (or the arc of the edge) which was traversed/served by the vehicle, it is represented as a tuple of the head node and tail node of the arc;
4. a bit value that indicates whether the arc was only traversed (0) or served (1), too.

For instance, "(25, 7, (2, 8), 1)" means the vehicle identified by 7, served the arc with head node 2 and tail node 8 in the 25-th minute. The time point is represented as the minutes passed since the execution of the first traveling/serving event (i.e., the first event was executed in the 0-th minute).

In this work, for the sake of simplicity, it is assumed that events are generated only when the vehicle starts traversing/serving an arc/edge and no dynamic events can occur that disturb the completion process of the currently executed activity. Therefore, a vehicle can breakdown only right after completing its current activity or right before starting the next one. Furthermore, an activity can only be correctly and fully completed (i.e., all of the demand on the arc/edge is served).

The travel and service log σ contains the events of the data stream as a sequence, in the order of observation.

3.2. "Change in problem" event stream

In this work, it is assumed that when a dynamic event occurs while the solution of the current problem is being executed, the necessary information about this event arrives through an event stream, the "change in problem" event stream. This information is different for every event, so there is not a general structure description. However, all the events must contain the time point when they occurred, so it can be determined that in which state the execution of the solution of the problem was (according to the travel and service log) when the change happened. It is important that the way of representation (the format) of the time point has to be the same as in the travel and service log. In the scope of this work, only three events are considered in greater detail: task appearance, demand increase and vehicle breakdown (for the reason and for a detailed description of the events, see Section 4).

For a task appearance event, the followings are expected:

1. the arc (or one of the arcs of an edge) $a \in A$ where the new task appeared;
2. demand of the new task;
3. service cost of the new task.

For a demand increase event, the following are expected:

1. the arc (or one of the arcs of an edge) $a \in A$ where the demand increased;
2. the increment in the demand of the task;
3. the increment in the service cost of the task;

For a vehicle breakdown event, only the identifier of the broken down vehicle is expected.

4. ANALYZING THE POSSIBLE EVENTS

Assumptions:

- The traveling cost of an edge/arc is determined by the length of the corresponding road segment (static value) and the state of the traffic there (dynamic value).
- The service cost of a task edge/arc is determined by the traveling cost and the cost of the subtasks within the task (dynamic value). The cost of the subtasks within a task is determined by the demand. Therefore, if the demand increases/decreases, the service cost increases/decreases, as well.
- The road segments with tasks cannot be closed (since it would make the problem unsolvable), and if a road segment without task gets closed, it does not change the reachability property of the graph (i.e., there is always a path between node A and B, if $A, B \in V$).

The possible events can be divided into two main groups, based on what type of component of the problem they are related to; the road segments (i.e., the graph) or the vehicles. In Section 4.1, the possible events related to road segments, and in Section 4.2, the possible events related to vehicles are analyzed. In Section 4.3, conclusions are made and the existing publications on D(C)ARP are compared based on which type of dynamic events they consider.

4.1. Analyzing the possible events related to road segments

If a road gets closed, the corresponding road segment(s) will not be traversable anymore. Since this situation is temporary, instead of removing the corresponding edge(s) and arc(s), only their traveling costs are changed to infinite value (*inf*). It does not make rerouting necessary, but it can make the route plans infeasible in the current solution, if the affected edge(s) and arc(s) are supposed to be traversed throughout according to at least one of the plans. In such cases, the shortest path between the affected tasks should be recalculated with the new cost values, then the route plans in question should be updated in the solution. In case the new shortest path between two tasks includes one or multiple unexecuted task(s) of the route plan, the task(s) in question can be inserted there (so the order in which the tasks are executed changes), then the shortest paths between the affected tasks should be recalculated again. The solution still might not become optimal, but it becomes feasible.

If a road gets opened, the corresponding road segment(s) will be traversable again so their traveling costs are changed (based on the length of the road and the state of traffic there) to new values. It does not make rerouting necessary and has no effect on the feasibility of the current solution. However, rerouting can improve the current solution, if using the opened road segments in the route plans decrease the total cost.

If the traffic decreases on a road or road segment, it decreases the traveling cost of the corresponding edge(s) and arc(s). Furthermore, if any of these edge(s) and arc(s) have unserved tasks, then their service cost decrease, too. It does not make rerouting necessary and has no effect on the feasibility of the current solution. However, rerouting can improve the current solution, if using the affected road segments in the route plans decreases the total cost.

If the traffic increases on a road or road segment, it increases the traveling cost of the corresponding edge(s) and arc(s). Furthermore, if any of these edge(s) and arc(s) have unserved tasks, then their service cost increases, too. It does not make rerouting necessary and has no effect on the feasibility of the current solution. However, rerouting can improve the current solution, if using other road segments instead of the affected road segments in the route plans decrease the total cost.

If a task is canceled (i.e., no subtasks left) on a road segment, the corresponding edge/arc is removed from the required edges/arcs. It does not make rerouting necessary, but makes the current solution obsolete. The canceled task should be removed from the corresponding route plan of the solution, and then it should be updated with the shortest path between the preceding and the following task of the removed task (according to the plan). Rerouting can further improve the updated solution, if serving the tasks in the modified route plan in another order decreases the total cost.

If a new task appears on a road segment that did not have any tasks before (or the tasks were already served beforehand), the corresponding edge/arc is added to the required edges/arcs with the defined service cost and demand. It makes rerouting necessary and the current solution obsolete, since in the current solution the new task is not planned to be served and it is not self-evident in which route plan should the new task be added to.

If the demand increases on a road segment, the demand also increases on the corresponding edge/arc task and the service cost may increase too. It makes rerouting necessary and the current solution infeasible, if the demand of the task increased in an extent that the total demand of the tasks in the route plan exceeds the capacity of the vehicle.

If the demand decreases on a road segment, the demand also decreases on the corresponding edge/arc task and the service cost may decrease too. It does not make rerouting necessary and has no effect on the feasibility of the current solution. However, rerouting can improve the current solution, if the task with its decreased demand now can be served by another vehicle with lower total cost.

Table 1. Summary of the possible events related to road segments and their effect(s) on the problem

Event	Effect on the problem	Modification	Rerouting
road closed	traveling cost *dc* of the affected edge(s) and arc(s) changes to *inf* (from *x*)	yes*	no
road opened	traveling cost *dc* of the affected edge(s) and arc(s) changes to *x* (from *inf*)	no	no
traffic decreased	decreases the traveling cost *dc* of the affected edge(s) and arc(s) (for unserved tasks, decreases the service cost *sc* too)	no	no
traffic increased	increases the traveling cost *dc* of the affected edge(s) and arc(s) (for unserved tasks, increases the service cost *sc* too)	no	no
task cancellation	removes the edge/arc from the required edges/arcs	yes	no
task appearance	adds the edge/arc to the required edges/arcs with demand *d* and service cost *sc*	yes	yes
demand increased	increases demand *d* of the edge/arc task by Δd and the service cost *sc* by Δsc	yes*	yes*
demand decreased	decreases demand *d* of the edge/arc task by Δd and the service cost *sc* by Δsc	no	no

The possible events related to road segments are summarized in Table 1. In the "Modification" column, it is indicated whether modification of the current solution is necessary. Likewise, in the "Rerouting" column, it is indicated whether rerouting is necessary. The asterisk means conditional answer (i.e., modification/rerouting is necessary only if the effect of the event on the solution/problem satisfies a condition).

4.2. Analyzing the possible events related to vehicles

If a vehicle breaks down, the number of available vehicles decreases by one. It makes rerouting necessary and the current solution infeasible, if the vehicle still has unserved tasks on its route plan.

If a broken down vehicle restarts, then the number of available vehicles increases by one. It does not make rerouting necessary and has no effect on the feasibility of the current solution. However, rerouting can improve the current solution, if the restarted vehicle can serve a subset of the unserved tasks with lower total cost than the other vehicles.

The possible events related to vehicles are summarized in Table 2. Just like in Table 1, in the "Modification" column, it is indicated whether modification of the current solution is necessary. Likewise, in the "Rerouting" column, it is indicated whether rerouting is necessary. The asterisk means conditional answer (i.e., modification/rerouting is necessary only if the effect of the event on the solution/problem satisfies a condition).

Table 2. Summary of the possible events related to vehicles and their effect(s) on the problem

Event	Effect on the problem	Modification	Rerouting
vehicle breakdown	the number of available vehicles decreases by one (if it has unserved tasks, they should be served by another vehicle(s))	yes*	yes*
vehicle restart	the number of available vehicles increases by one	no	no

4.3. Conclusions

In this section, the possible events were collected and analyzed in the scope of their effect on the problem. For each event, it was examined whether modification of the current solution and rerouting is necessary in case of their occurrence to obtain a feasible solution. In a few cases (i.e., road closure and task cancellation events), updating of the current solution can be easily achieved with simple algorithms so rerouting is not necessary. In another few cases (i.e., task appearance, demand increased, and vehicle breakdown events), updating of the current solution requires a more complex algorithm, so rerouting is necessary. From now on, these events are collectively called critical events.

In the table below (Table 3), the possible events and literature on D(C)ARPs are shown. The content of the table is mainly based on the literature review on DARPs in [13]. The numbers in the header are references to the publications. The marks (X) indicate that the referred publication considers that event. The asterisk on the right side of the name of an event indicates that the event might be named differently in the referred publications, but nonetheless has the same effect on the problem. The critical events are indicated in bold, and events that none of the publications considers are colored gray. It can be seen, that only the most recent publication on DCARP [23] considers all the critical events. Furthermore, it is the publication that considers most of the possible events.

Table 3. The possible events and literature on D(C)ARPs

Event	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]
road closed		X	X					X	X
road opened									X
traffic improved*									X
traffic worsened*		X					X		X
task cancellation	X	X		X					
task appearance	**X**	**X**			**X**	**X**			**X**
demand increased		**X**	**X**						**X**
demand decreased									
vehicle breakdown								**X**	**X**
vehicle restart									

In the scope of this work, algorithms were developed to reroute the route plans of the current solution when events occur that make it necessary. These algorithms are introduced in the next section (Section 5).

5. DEVELOPED ALGORITHMS

In this section, three event handling algorithms are introduced that use the virtual task strategy [23] and RR1, a novel rerouting algorithm that reroutes at most one route plan of the current solution, while aiming to find the least cost route plans. The algorithms use the travel and service log of the vehicles, since there can be differences between the route plans and the execution of these plans.

The algorithms that handle the task appearance event, the demand increased event and the vehicle breakdown event are described in Section 5.1, Section 5.2 and Section 5.3, respectively. The algorithm that updates the virtual tasks in the current instance and solution is described in Section 5.4 and the RR1 rerouting algorithm is described in Section 5.5. For the sake of simplicity, error handlings are omitted from the algorithmic descriptions.

5.1. Handling the Task Appearance Event

The algorithm that handles the task appearance event can be seen in Fig. 1. The inputs of the algorithm are the current instance I, the current solution S, the travel and service log σ, the index of the last processed event in the log σ (i_σ), and the information about the new task: the arc/edge where the task appeared (a), the demand of the new task (dem_a) and the service cost of it (sc_a), respectively.

First, the served tasks are removed from the set of tasks T (line 2) and the virtual tasks are updated in both the current instance I and the solution S (line 3), then the new task is added to the given arc (lines 4-6). If the arc a is actually one of the arcs of an edge (i.e., the given arc a has an inverse arc $inv(a)$, so the arc a is an element of the domain of the inv function, $a \in dom(inv)$), then the task is added to the inverse arc $inv(a)$, too (lines 7-10). S_1 is a solution that does not contain the new task and S_2 is a solution that contains the new task in a new route (line 11).

The *RR1* algorithm tries to insert the new task into one of the current routes of S_1 while keeping the route change cost (i.e., the change in the total cost of the solution) minimal (line 12). If none of the outside vehicles have enough capacity or adding the task to an existing route is more expensive than adding it to a new route, then S_2 will be the new solution S. Finally, the algorithm returns the updated instance and the new solution (line 13).

Algorithm 1: Handling a "Task Appearance" Event

input : $I = (V, v_0, A, T, T_v, Q, H, H_f, R, R_e, rt, rv, head, tail, dc, inv, dem, sc, mdc)$,
$S = \{r_1, r_2, ..., r_K\}, \sigma \in (\mathbb{N}^+ \times \mathbb{N}^+_{\leq |H|} \times A \times \{0, 1\})^*, i_\sigma \in \mathbb{N}^+$,
$a \in A, dem_a \in \mathbb{R}^+, sc_a \in \mathbb{R}^+$

1 **begin**
2 $I \leftarrow removeServedTasks(I, \sigma, i_\sigma)$;
3 $I, S_1, i_\sigma \leftarrow updateVirtualTasks(I, S, \sigma, i_\sigma, -1)$;
4 $dem(a) \leftarrow dem_a$;
5 $sc(a) \leftarrow sc_a$;
6 $T \leftarrow T \cup a$;
7 **if** $a \in dom(inv)$ **then**
8 $dem(inv(a)) \leftarrow dem_a$;
9 $sc(inv(a)) \leftarrow sc_a$;
10 $T \leftarrow T \cup inv(a)$;
11 $S_2 \leftarrow S_1 \cup \{\langle v_0, a, v_0 \rangle\}$;
12 $I, S \leftarrow RR1(I, S_1, S_2, \langle a \rangle)$;
13 **return** I, S;

Fig. 1. Algorithm for handling the task appearance event

5.2. Handling the Demand Increased Event

The algorithms that handles the demand increased event can be seen in Fig. 2. The inputs of the algorithm are the current instance I, the current solution S, the travel and service log σ, the index of the last processed event in the log σ (i_σ), and the information about the task in which the demand increased: the arc/edge where the task is (a), the increment in the demand of the task (Δdem_a) and the increment in the service cost of it (Δsc_a), respectively.

First, the served tasks are removed from the set of tasks T (line 2) and the virtual tasks are updated in both the current instance I and the solution S (line 3), then the current demand $dem(a)$ and the service cost $sc(a)$ values of the given arc task are increased by the given increment values (lines 4-5). If the arc task a is actually one of the arcs of an edge task (i.e., the given arc a has an inverse arc inv(a)), then the demand and the service cost values of the inverse arc task are increased, too (lines 6-8).

Next, it is checked whether the vehicle still has enough capacity to serve all the remaining tasks on its route where the task with increased demand is included (lines 9-15). For this, the index of the route plan (i) in which the task is, has to be found (lines 9-12), then the load of the vehicle (i.e., the sum of the demands of all the tasks on the route r_i) is calculated (line 13). Assuming that the vehicle starts empty in case of collection service (e.g., urban waste collection), and full in case of distribution service (e.g., salting the streets), if the load does not exceed the maximum capacity Q of the vehicle, then the solution S_1 is still feasible, so it can be returned with the updated instance I (lines 14-15). Otherwise, rerouting is needed. The task with increased demand a is removed from the solution S_1 (line 16), then added to a new route and recorded as solution S_2 (line 17). Therefore, S_1 is a solution that does not contain the task a and S_2 is a solution that contains the task a in a new route.

The same way as in the task appearance event case (Section 5.1), the *RR1* algorithm tries to insert the task into one of the current routes of S_1 while keeping the route change cost (i.e., the change in the total cost of the solution) minimal (line 18). If none of the outside vehicles have enough capacity or adding the task to an existing route is more expensive than adding it to a new route, then S_2 will be the new solution S. Finally, the algorithm returns the updated instance and the new solution (line 19).

Algorithm 2: Handling a "Demand Increased" Event

input : $I = (V, v_0, A, T, T_v, Q, H, H_f, R, R_e, rt, rv, head, tail, dc, inv, dem, sc, mdc)$,
$S = \{r_1, r_2, ..., r_K\}, \sigma \in (\mathbb{N}^+ \times \mathbb{N}^+_{\leq |H|} \times A \times \{0, 1\})^*, i_\sigma \in \mathbb{N}^+$,
$a \in A, \Delta dem_a \in \mathbb{R}^+, \Delta sc_a \in \mathbb{R}^+$

1 **begin**
2 $I \leftarrow removeServedTasks(I, \sigma, i_\sigma)$;
3 $I, S_1, i_\sigma \leftarrow updateVirtualTasks(I, S, \sigma, i_\sigma, -1)$;
4 $dem(a) \leftarrow dem(a) + \Delta dem_a$;
5 $sc(a) \leftarrow sc(a) + \Delta sc_a$;
6 **if** $a \in dom(inv)$ **then**
7 $dem(inv(a)) \leftarrow dem(a) + \Delta dem_a$;
8 $sc(inv(a)) \leftarrow sc(a) + \Delta sc_a$;
9 $i \leftarrow -1$;
10 **forall** $r_j \in S_1$ **do**
11 **if** $a \in r_j$ **then**
12 $i \leftarrow j$;
13 $load_i \leftarrow load(I, r_i)$;
14 **if** $load_i \leq Q$ **then**
15 **return** I, S_1;
16 $S_1 \leftarrow removeTask(S_1, a)$;
17 $S_2 \leftarrow S_1 \cup \langle v_0, a, v_0 \rangle$;
18 $I, S \leftarrow RR1(I, S_1, S_2, \langle a \rangle)$;
19 **return** I, S;

Fig. 2. Algorithm for handling the demand increased event

5.3. Handling the Vehicle Breakdown Event

The algorithm that handles the vehicle breakdown event can be seen in Fig. 3. The inputs of the algorithm are the current instance I, the current solution S, the travel and service log σ, the index of the last processed event in the log σ (i_σ), and the identifier of the broken down vehicle (i), respectively.

First, the served tasks are removed from the set of tasks T (line 2), the virtual tasks are updated in both the current instance I and the solution S (line 3), and the identifier of the broken down vehicle is added to the set of IDs that cannot serve tasks R_e (line 4). In this case, the *updateVirtualTask* function returns two solutions, S_1 and S_2, and a list of tasks T_i that were included in the route plan of the broken down vehicle i and are not served yet. The solution S_1 does not contain the tasks of T_i and S_2 contains them in a new route. If the broken down vehicle does not have any

unserved tasks (i.e., the size of T_i is 0) (line 5), then S_1 is returned as the new solution with the updated instance I (line 6). Otherwise, rerouting is needed.

The goal of the *RR1* algorithm is to minimize the number of modified route plans while keeping the total cost minimal. Therefore, it tries to insert the tasks of T_i into one of the current route plans of S_1 while keeping the route change cost (i.e., the change in the total cost of the solution) minimal (line 8). If none of the outside vehicles have enough capacity or adding the tasks to an existing route is more expensive than adding them to a new route, then S_2 will be the new solution S. Finally, the algorithm returns the updated instance and the new solution (line 9).

Algorithm 3: Handling a "Vehicle Breakdown" Event

input : $I = (V, v_0, A, T, T_v, Q, H, H_f, R, R_e, rt, rv, head, tail, dc, inv, dem, sc, mdc)$,
$S = \{r_1, r_2, ..., r_K\}, \sigma \in (\mathbb{N}^+ \times \mathbb{N}^+_{\leq |H|} \times A \times \{0,1\})^*, i_\sigma \in \mathbb{N}^+, i \in \mathbb{N}^+_{\leq |H|}$

1 **begin**
2 $I \leftarrow removeServedTasks(I, \sigma, i_\sigma)$;
3 $I, S_1, S_2, T_i, i_\sigma \leftarrow updateVirtualTasks(I, S, \sigma, i_\sigma, -1)$;
4 $R_e \leftarrow R_e \cup i$;
5 **if** $|T_i| = 0$ **then**
6 **return** I, S_1;
7 **else**
8 $I, S \leftarrow RR1(I, S_1, S_2, T_i)$;
9 **return** I, S;

Fig. 3. Algorithm for handling the vehicle breakdown event

5.4. Updating Virtual Tasks

The algorithm for adding and updating virtual tasks *updateVirtualTasks* can be seen in Fig. 4. The construction process of the virtual tasks is mainly based on the "Pseudocode of constructing virtual tasks" (Algorithm 2) in [23], but there are some differences which are described in more detail in Section 2.2.

The inputs of the algorithm are the current instance I, the current solution S, the travel and service log σ, the index of the last processed event in the log σ (i_σ), and the identifier of the broken down vehicle i_e (only in the case of vehicle breakdown event, otherwise -1 is given), respectively.

First, the solutions S_1 and S_2, and the sequence T_e that will store the unserved tasks of the broken down vehicle are initialized (lines 2-4), then the iteration starts in the route plans of the current solution to construct and add virtual tasks to them or update the previous ones (lines 6-50). (Note: S_2 and T_e are used only in the case of vehicle breakdown event.)

First, the identifier of the vehicle i_k that executes the route plan r_k is obtained with the *rv* function (line 6). Next, the variable v_k that will store the current position of the vehicle is initialized (line 7), along with the updated version of the route (r), the total cost (*tc*) and total demand (*tdem*) variables (line 8-10), that will store the change in the total cost and demand of the route r_k, respectively. As it was mentioned before, it is assumed that the current position of the vehicle is on a vertex.

Next, by iterating over the new events in the travel and service log σ (i.e., the events that are not processed yet, based on i_σ), the new events that were executed by the vehicle are obtained and examined one-by-one (line 11-24). In each step, the current position of the vehicle (v_k) is updated (line 12), then it is checked whether the vehicle finished its current route (i.e., reached the depot) (line 13). In such a case, the identifier of the route (k) is added to the set R_e (which contains the identifiers of the routes that are finished or got suspended) (line 14). If there are any route plans for which no vehicle was assigned yet (line 15), then the identifier of the vehicle (i_k) is assigned to the oldest one of them (i.e., to the route plan with the smallest identifier) with the function *rv* (lines 16-17). Otherwise, the identifier of the vehicle (i_k) is added to the set that contains the identifiers of the free vehicles (line 19). Regardless of the current position of the vehicle, if the vehicle only traveled throughout arc a (i.e., $b = 0$), then only the traveling cost $dc(a)$ is added to the total cost *tc* (lines 20-21). If the vehicle served arc a (i.e., $b = 1$), then the service cost $sc(a)$ is added to the total cost *tc* and the demand $dem(a)$ is added to the total demand *tdem* (lines 22-24). Here it is assumed that one arc/edge can be served by only one vehicle (i.e., a task cannot be divided into subtasks and be executed partially).
Next, the (new) virtual task *vt* is constructed, but only if the vehicle moved since the last time the virtual tasks were updated (i.e., the variable v_k does not have its initial value, so there is at least one new event generated by the vehicle) (line 25-39). There are two cases: (1) there is already a virtual task assigned to the route plan r_k and (2) there is no virtual task assigned to the route plan yet (since it was started to be executed only recently). In both cases, the (new) virtual task *vt* is constructed the same way (line 26) and its *head*, *tail*, and *dc* are defined the same way as well (lines

27-29). However, in case (1), the service cost (*sc*) is calculated as the sum of the service cost of the previous virtual task (*sc*(*rt*(*k*))) and the change in the total cost since then (*tc*) (line 31). The demand is calculated likewise (line 32),

Algorithm 4: updateVirtualTasks

input : $I = (V, v_0, A, T, T_v, Q, H, H_f, R, R_e, rt, rv, head, tail, dc, inv, dem, sc, mdc)$,
$S = \{r_1, r_2, ..., r_K\}, \sigma \in (\mathbb{N}^+ \times \mathbb{N}^+_{\leq |H|} \times A \times \{0,1\})^*, i_\sigma \in \mathbb{N}^+, i_e \in \{\mathbb{N}^+_{\leq |H|}, -1\}$

1 **begin**
2 $S_1 \leftarrow \emptyset$;
3 $S_2 \leftarrow \emptyset$;
4 $T_e \leftarrow \langle\rangle$;
5 **forall** $r_k \in S$ **do**
6 $i_k \leftarrow rv(r_k)$;
7 $v_k \leftarrow -1$;
8 $r \leftarrow \langle v_0 \rangle$;
9 $tc = 0$;
10 $tdem = 0$;
11 **forall** $e_m = (t, i, a, b) \in \sigma$ *s.t.* $i = i_k \wedge m > i_\sigma$ **do**
12 $v_k \leftarrow tail(a)$;
13 **if** $v_k = v_0$ **then**
14 $R_e \leftarrow R_e \cup \{k\}$;
15 **if** $\exists x \in R$ *s.t.* $rv(x) = -1$ **then**
16 $x \leftarrow min\{x | x \in R, rv(x) = -1\}$;
17 $rv(x) \leftarrow i_k$;
18 **else**
19 $H_f \leftarrow H_f \cup \{i_k\}$;
20 **if** $b = 0$ **then**
21 $tc \leftarrow tc + dc(a)$;
22 **if** $b = 1$ **then**
23 $tc \leftarrow tc + sc(a)$;
24 $tdem \leftarrow tdem + dem(a)$;
25 **if** $v_k \neq -1$ **then**
26 $vt \leftarrow (v_0, v_k)$;
27 $head(vt) \leftarrow v_0$;
28 $tail(vt) \leftarrow v_k$;
29 $dc(vt) \leftarrow \infty$;
30 **if** $rt(k) \neq -1$ **then**
31 $sc(vt) \leftarrow tc(rt(k)) + tc$;
32 $dem(vt) \leftarrow dem(rt(k)) + tdem$;
33 $T_v \leftarrow \{T_v \setminus rt(k)\} \cup vt$;
34 **else**
35 $sc(vt) \leftarrow tc$;
36 $dem(vt) \leftarrow tdem$;
37 $T_v \leftarrow T_v \cup vt$;
38 $rv(k) \leftarrow vt$;
39 $r \leftarrow r \cdot \langle vt \rangle$;
40 **if** $v_k = -1 \wedge rt(k) \neq -1$ **then**
41 $r \leftarrow r \cdot \langle rt(k) \rangle$;
42 **forall** $a \in r_k$ **do**
43 **if** $a \in T$ **then**
44 $r \leftarrow r \cdot \langle a \rangle$;
45 **if** $i_k = i_e$ **then**
46 $T_e \leftarrow T_e \cdot \langle a \rangle$;
47 $r_k \leftarrow r \cdot \langle v_0 \rangle$;
48 $S_1 \leftarrow S_1 \cup \{r_k\}$;
49 **if** $i_k \neq i_e$ **then**
50 $S_2 \leftarrow S_2 \cup \{r_k\}$;
51 $i_\sigma \leftarrow |\sigma|$;
52 **if** $i_e \neq -1$ **then**
53 **return** $I, S_2, S_1, T_e, i_\sigma$;
54 **else**
55 **return** I, S_1, i_σ;

Fig. 4. Algorithm for adding virtual tasks

then the old virtual task is removed and the new one is added to the set of virtual tasks T_v (line 33). In case (2), the service cost is equal to change in the total cost (line 35) and the demand to the change in the total demand (line 36). Since it is the first virtual task of the route, it is added to the set of virtual tasks T_v without removing another one (line 37). As final steps (in both cases), the new virtual task is assigned to the identifier of the route plan (line 38), then added to the updated route plan *r* (line 39).

If the vehicle did not move, but it has a virtual task (line 40), then its current virtual task is the same as the previous one. Therefore, the previous virtual task *rt*(*k*) is added to the updated route plan *r* (line 41).

After the virtual task is constructed and added to the route plan *r*, the unserved tasks of the previous route plan are added one-by-one (lines 42-44) finished by the dummy task v_0 (line 47), then added into the new solution S_1 (line 48). In the case of vehicle breakdown event, the unserved tasks of the broken down vehicle are collected into T_e (lines 45-46). Furthermore, a solution S_2 is created too, which does not contain the unserved tasks of the broken down vehicle (S_1 contains them in a new route) (lines 49-50).

Finally, the algorithm updates the index of the last processed event in the log σ (i.e., i_σ) (line 51), then returns the updated instance *I*, the new solution S_1, and the updated log index i_σ (lines 54-55). In case of vehicle breakdown event, solution S_2 and the sequence of the unserved tasks of the broken down vehicle T_e are returned, too (lines 52-53).

5.5. RR1 rerouting algorithm

The algorithm of rerouting *RR1* can be seen in Fig. 5. The inputs of the algorithm are the current instance *I*, the candidate solutions S_1 and S_2, and the sequence of tasks T_r that needs to be inserted into solution S_2. The solution S_1 contains all the tasks from T_r in a separate route plan and the solution S_2 does not contain any of the tasks from T_r, hence this algorithm tries to insert them into S_2 the most cost effective way.

First, the total demand of the tasks in T_r is calculated (lines 2-4), then the identifiers of the route plans are collected into the set *ids* that have enough remaining capacity to serve all the tasks in T_r (lines 5-8). The load on a route of a vehicle (i.e., the sum of the demands of the served tasks and the tasks that are planned to be served) is calculated with the *load* function from the corresponding route plan in S_2. Just like in Section 5.2, it is assumed that a vehicle starts empty in case of collection service (e.g., urban waste collection), and full in case of distribution service (e.g., salting the streets). If there is no route plan with enough capacity (i.e., *ids* is an empty set), then the tasks can only be added to a new route plan, hence S_1 is returned as the new solution (lines 9 and 50). The tasks can certainly fit into one new route plan, since it is assumed that only one vehicle can break down at a time (i.e., in one event) and that the vehicles are homogenous (i.e., they have the same maximum capacity). Otherwise, for each vehicle *i* in *ids*, the algorithm tries to insert the tasks of T_r one-by-one into the position that generates the less extra cost (lines 11-39).

First, the variables that store the best solution *S* and best total cost *tc* so far are initialized with solutions S_1 (line 9) and its total cost (line 11), then the iteration starts in the *ids*. For each vehicle *i* in *ids*, the best solution is temporarily stored in S_i and the total cost of it is tc_i. In each iteration step, these variables are initialized with solution S_2 and the total cost of it (lines 13-14), then the iteration over the tasks of T_r starts. For each task *a* in T_r, the best position is searched within the route plan r_i (lines 15-39). For this searching process, the smallest total cost change Δtc_i of the insertion of task *a* within route plan r_i, the position of this insertion j_a, and whether the task *a* should be inverted inv_a are stored temporarily. These variables are initialized with the values infinity, -1 and False, respectively (lines 16-18). Next, the iteration of the possible positions within the route plan r_i starts (lines 19-31). Task *a* can be inserted only between the dummy tasks and only after the virtual task. For each position *j*, the total cost change of insertion of task *a* to position *j* within the route plan r_i (i.e., Δtc_{i1}) is calculated (line 20). If task *a* is an edge task, then the total cost change of insertion of *inv(a)* (i.e., Δtc_{i2}) is calculated as well (lines 22-23). If *j* is the best position to insert task *a* within the route plan r_i so far (i.e., with *j* smaller total cost change can be realized than with j_a), then it will be the new best position (line 24-27). The same applies to *inv*(*a*) (lines 28-31). After all the positions are examined, task *a* (or *inv*(*a*) if inv_a = True) is inserted with the function *insert* to the known best position j_a within the route plan r_i of the solution S_i (lines 32-35). The total cost tc_i of the solution S_i is also increased by the smallest known total cost change Δtc_i (i.e., the total cost change after the insertion of the task) (line 36). After all the tasks in T_r are inserted into solution S_i, the algorithm checks if this solution is better (i.e., has smaller total cost) than the currently known best solution *S* (line 37). If the solution S_i is better, then it becomes the new best solution and the total cost of it becomes the new best total cost (lines 38-39).

After all the route plans are examined, if the best solution is to put all the tasks into a new route (i.e., S_1) (line 40), then the identifier for the new route is created (line 41) and added to the set of route identifiers *R* (line 42). Next, its virtual task *rt* is initialized with its default value (line 43). If there are any free vehicles available (line 44), then one of them is assigned to the new route plan (lines 45-47), otherwise *rv* is initialized with its default value (line 49).

After all the route plans are examined and (in case of new route insertion) the instance is updated, the updated instance *I* and the best solution *S* are returned by the algorithm (line 50).

Algorithm 5: RR1

input : $I = (V, v_0, A, T, T_v, Q, H, H_f, R, R_e, rt, rv, head, tail, dc, inv, dem, sc, mdc)$, $S_1 = \{r_{11}, r_{12}, ..., r_{1K_1}\}$, $S_2 = \{r_{21}, r_{22}, ..., r_{2K_2}\}$, $T_r \in T^*$

1 **begin**
2 $dems \leftarrow 0$;
3 **forall** $a \in T_r$ **do**
4 $dems \leftarrow dems + dem(a)$;
5 $ids \leftarrow \emptyset$;
6 **forall** $i \in \{1, 2, ..., K_2\} \setminus R_e$ **do**
7 **if** $load(I, r_{2i}) + dems \leq Q$ **then**
8 $ids \leftarrow ids \cup \{i\}$;
9 $S \leftarrow S_1$;
10 **if** $ids \neq \emptyset$ **then**
11 $tc \leftarrow TC(I, S_1)$;
12 **forall** $i \in ids$ **do**
13 $S_i \leftarrow S_2$;
14 $tc_i \leftarrow TC(I, S_2)$;
15 **forall** $a \in T_r$ **do**
16 $\Delta tc_i \leftarrow \infty$;
17 $j_a \leftarrow -1$;
18 $inv_a \leftarrow False$;
19 **forall** $t_{i,j} \in r_i$ *s.t.* $t_{i,j} \notin \{v_0\} \cup T_v$ **do**
20 $\Delta tc_{i1} \leftarrow sc(a) + mdc(tail(t_{i,j-1}), head(a)) + mdc(tail(a), head(t_{i,j})) - mdc(tail(t_{i,j-1}), head(t_{i,j}))$;
21 $\Delta tc_{i2} \leftarrow \infty$;
22 **if** $a \in dom(inv)$ **then**
23 $\Delta tc_{i2} \leftarrow sc(inv(a)) + mdc(tail(t_{i,j-1}), head(inv(a))) + mdc(tail(inv(a)), head(t_{i,j})) - mdc(tail(t_{i,j-1}), head(t_{i,j}))$;
24 **if** $\Delta tc_{i1} < \Delta tc_i$ **then**
25 $\Delta tc_i \leftarrow \Delta tc_{i1}$;
26 $j_a \leftarrow j$;
27 $inv_a \leftarrow False$;
28 **if** $\Delta tc_{i2} < \Delta tc_i$ **then**
29 $\Delta tc_i \leftarrow \Delta tc_{i2}$;
30 $j_a \leftarrow j$;
31 $inv_a \leftarrow True$;
32 $t \leftarrow a$;
33 **if** $inv_a = True$ **then**
34 $t \leftarrow inv(a)$;
35 $S_i \leftarrow insert(S_i, i, j_a, t)$;
36 $tc_i \leftarrow tc_i + \Delta tc_i$;
37 **if** $tc_i \leq tc$ **then**
38 $tc \leftarrow tc_i$;
39 $S \leftarrow S_i$;
40 **if** $S = S_1$ **then**
41 $k \leftarrow |R|$;
42 $R \leftarrow R \cup \{k\}$;
43 $rt(k) \leftarrow -1$;
44 **if** $H_f \neq \emptyset$ **then**
45 $i \leftarrow randomChoice(H_f)$;
46 $rv(k) \leftarrow i$;
47 $H_f \leftarrow H_f \setminus i$;
48 **else**
49 $rv(k) \leftarrow -1$;
50 **return** I, S

Fig. 5. The RR1 rerouting algorithm

6. GENERATING EVENTS

To be able to test the developed DCARP framework, the event handling algorithms, and the RR1 rerouting algorithm, travel and service logs with various content and length, and events with various event data have to be used. For this reason generators were developed for the travel and service log and the discussed three events. The travel and service log generator constructs a log with random length, based on the given initial DCARP instance and its solution. Since the initial DCARP is a static CARP, the initial solution can be obtained by using solvers for static CARP (e.g., [7], [9]). The event generators attempt to create events randomly, based on the log and the given DCARP instance.

6.1. Generating the travel and service log

As it was described in Section 3.1, the travel and service log is assumed to be a sequence of events, where an event contains information about the time point when the vehicle started to travel/serve the arc, the vehicle (the unique identifier of the vehicle), the traversed/served arc, whether the arc was only traversed (0) or served (1), too (by the vehicle).

The events in the log are generated by simulating the execution of the given solution sequence *S*. Since *S* contains only the served arcs, the traversed arcs are added based on the shortest paths between the served arcs. Assuming that the traversing cost and the service cost of the arcs are defined by the time spent on serving/traversing them, the total cost so far is used as the time point of the events and also in the stopping condition of the execution of each route. To provide diversity in the generated logs, the total cost where the execution of a route plan stops (*crc*) is chosen randomly in the range 0 - *max_route_cost*, where *max_route_cost* is the total cost of the most expensive (i.e., the most time consuming) route plan among the route plans of the given solution *S*. Therefore, if the value of *crc* is 0, then the log will be empty, since the execution of the route plans is not started yet. Likewise, if the value of *crc* is *max_route_cost*, the log will be complete, since the execution of all the route plans is finished.

The algorithm of generating the travel and service log works as follows:

1. Get the cost of the most expensive route plan among the route plans of the given solution *S*. → *max_route_cost*
2. Choose a number between 0 and *max_route_cost* randomly. → *crc*
3. Iterate over the route plans of the solution *S* (in parallel) and add the served/traveled arc to the log only if the sum of the service/traversing cost of the arc and the total cost so far of the route does not exceed *crc*. → *log*
4. Return *log*.

6.2. Generating data for the task appearance event

As it was described in Section 3.2, for the task appearance event, the arc (where the task appeared), the demand of the task, and the service cost of the task are recorded. The arc can only be a currently unrequired arc, so the arc is either unrequired since the beginning, or it is a previously required arc, that has already been served (based on the travel and service log).

To provide diversity in the generated events, the arc (*nt_arc*) is chosen randomly from the currently unrequired arcs (*unrequired_arcs*). The demand (*nt_dem*) is also chosen randomly in the range $1 – Q$, since the demand has to be greater than 0 and it cannot exceed the maximum capacity of the vehicles (otherwise, it would be impossible to serve the task, assuming that a task can be served by only one vehicle). For the sake of simplicity, the service cost of the task (*nt_sc*) is defined to be the same as the traversing cost of the arc. The time point (*nt_time_point*) of the occurrence of the event is defined to be the same as the time point of the (currently) last event in the travel and service log.

If there are not any currently unrequired arcs, the task appearance event cannot happen, so no such event is generated by the generator.

The algorithm of generating data for the task appearance event works as follows:

1. Get the currently unrequired arcs (including the arcs of the edges), the unrequired arcs and required (task) arcs that have already been served (based on the log). → *unrequired_arcs*
2. If *unrequired_arcs* is empty, then return.
3. Choose the time point of the last event in the travel and service log as the time point of the occurrence of this event. → *nt_time_point*

4. Choose an arc from the *unrequired_arcs* randomly. → *nt_arc*
5. Choose a number between 1 and *Q* (i.e., the maximum vehicle capacity) as the demand of the new task. → *nt_dem*
6. Choose the value of the service cost of the new task as the traversing cost of *nt_arc*. → *nt_sc*
7. Return *nt_time_point*, *nt_arc*, *nt_dem* and *nt_sc*.

6.3. Generating data for the demand increased event

As it was described in Section 3.2, for the demand increased event, the arc (where the demand increased), the increment in the demand of the task, and the increment in the service cost of the task are recorded. The arc can only be a currently required arc, so the arc is either required since the beginning, or it is a previously unrequired arc, that got task assigned by a task appearance event and is not served yet (based on the travel and service log).

To provide diversity in the generated events, the arc (*dit_arc*) is chosen randomly from the currently required arcs (*required_arcs*). Assuming that an increase in the demand increases the service cost of the task in the same rate, the increment in the demand (*dit_dem_inc*) and the increment in the service cost (*dit_sc_inc*) is calculated using the same percentage (*increase_percentage*). The value of *increase_percentage* is chosen randomly in such a way, that it has to increase the demand at least by 1 (otherwise, there is no increase in the demand) and it will not make the new value of the demand exceed the maximum capacity of the vehicles *Q*. The time point (*dit_time_point*) of the occurrence of the event is defined to be the same as the time point of the (currently) last event in the travel and service log.

If there are not any currently required arcs, the demand increased event cannot happen, so no such event is generated by the generator.

The algorithm that generates data for the demand increased event works as follows:

1. Get the currently required arcs (including the arcs of the edges), the arcs that have not been served yet (based on the log). → *required_arcs*
2. If *required_arcs* is empty, then return.
3. Choose the time point of the last event in the travel and service log as the time point of the occurrence of this event. → *dit_time_point*
4. Choose an unserved task arc from *required_arcs* randomly. → *dit_arc*
5. Choose an increase percentage for the demand and service cost of the chosen task randomly in such a way, that it will increase the demand at least by 1 and it will not make the new value of the demand exceed *Q* (maximum vehicle capacity). → *increase_percentage*
6. Calculate the increase in the demand of the task using *increase_percentage*. → *dit_dem_inc*
7. Calculate the increase in the serving cost of the task using *increase_percentage*. → *dit_sc_inc*
8. Return *dit_time_point*, *dit_arc*, *dit_dem_inc* and *dit_sc_inc*.

6.4. Generating data for the vehicle breakdown event

As it was described in Section 3.2, for the vehicle breakdown event, only the identifier (ID) of the broken down vehicle is recorded. A vehicle can break down only if it is still outside on the road (i.e., not in the depot).

To provide diversity in the generated events, the identifier of the broken down vehicle (*vb_id*) is chosen randomly from the identifiers of the outside vehicles (*outside_vehicles*). The time point (*vb_time_point*) of the occurrence of the event is defined to be the same as the time point of the (currently) last event in the travel and service log.

If there are not any vehicles currently outside, the vehicle breakdown event cannot happen, so no such event is generated by the generator.

The algorithm that generates data for the vehicle breakdown event works as follows:

1. Get the set of IDs of the vehicles that are currently outside (based on the log). → *outside_vehicles*

2. If *outside_vehicles* is empty, then return.
3. Choose the time point of the last event in the travel and service log as the time point of the occurrence of this event. → *vb_time_point*
4. Choose a vehicle ID from *outside_vehicles* randomly. → *vb_id*
5. Return *vb_time_point*, *vb_id*.

7. EXPERIMENTS

The proposed approach was implemented in Python 3.7 and tested in version 3.3.6 of Spyder development environment. The tests were performed on a laptop PC with Windows 10 operation system, equipped with an Intel(R) Core(TM) i5-3320M 2.60 GHz 2-core CPU and 8 GB of RAM.

7.1. Experiment Setup

The implemented solution was tested on the real-life based *EGL* benchmark test set of static CARP instances [3]. The instances originate from the data of a winter gritting application in Lancashire (UK). To obtain the solutions for the initial DCARP instances (which are fundamentally static CARP instances), the hybrid metaheuristic approach (HMA) [9] was used. Furthermore, for the comparison of a complete rerouting algorithm with a minimal rerouting algorithm like RR1, HMA was used as well. The travel and service logs and the events were generated with the algorithms introduced in Section 6.

The testing process for each CARP instance was carried out as follows:

1. Obtain solution S_0 for instance I_0 using HMA.
2. Generate a travel and service log σ from S_0 (Section 6.1).
3. Generate a task appearance event based on σ (Section 6.2), then call the event handler for task appearance events (Section 5.1). The event handler returns the new instance I^a and the solution S^a_{RR1} for the new instance, which was calculated by RR1 algorithm (Section 5.5).
4. Generate a demand-increased event based on σ (Section 6.3), then call the event handler for demand increased events (Section 5.2). The event handler returns the new instance I^b and the solution S^b_{RR1} for the new instance, which was calculated by RR1 algorithm (Section 5.5).
5. Generate a vehicle breakdown event based on σ (Section 6.4), then call the event handler for vehicle breakdown events (Section 5.3). The event handler returns the new instance I^c and the solution S^c_{RR1} for the new instance, which was calculated by RR1 algorithm (Section 5.5).
6. For each instance created by the event handling algorithms (i.e., I^a, I^b and I^c), calculate the solutions with the HMA as well.
7. Repeat from Step 2 the given number of times.

7.2. Result Analysis

The results of the tests were quite varied. For better representativeness, the results can be seen for the egl-e1-A instance in Table 4, Table 5, and Table 6, for the task appearance events, the demand increase events, and the vehicle breakdown events, respectively. In all the three tables, the first few columns contain the data that make it possible to reconstruct the inputs: the travel and service log (with *crc*), the task appearance event (with *nt_arc*, *nt_dem*, and *nt_sc*), the demand increase event (with *dit_arc*, *dit_dem_inc*, and *dit_sc_inc*), and the vehicle breakdown event (*vb_id*). The "RR1" column contains the best total cost calculated by RR1 algorithm. Likewise, the "HMA" column contains the best total cost calculated by the HMA algorithm. The lowest (best) values are indicated in bold. The "diff." column contains the difference between the best total cost calculated by the RR1 and the HMA algorithms (i.e., RR1 - HMA). The best total cost of the initial solution is 3548. It can be seen that for all the three events the results are quite varied.

For the task appearance events and the vehicle breakdown events, the RR1 algorithm gave competitive results, since it gave the same quality result in half of the cases (in a few cases it even gave better quality results) than the HMA algorithm. Moreover, the RR1 algorithm guaranteed to remove only the necessary tasks from the problematic

route plan to add them to at most one route plan, while the HMA could remove any kinds of tasks from any route plans, and then add them to different route plans.

The implemented RR1 algorithm has a much faster calculation time, than the HMA. However, it must be noted, that the original implemented version of the HMA was not available, so it had to be implemented based on its algorithmic description (in [9]). Therefore, it is likely to have errors that decrease its effectiveness.

Table 4. The best total costs calculated by RR1 and HMA after task appearance events in egl-e1-A

crc	*nt_arc*	*nt_dem*	*nt_sc*	RR1	HMA	diff.
315	(25, 75)	58	16	4113	**3989**	124
315	(46, 45)	67	12	3911	**3900**	11
356	(43, 42)	11	11	3720	**3642**	78
379	(42, 57)	20	14	3698	**3666**	32
406	(2, 1)	62	32	**3612**	3628	-16
409	(73, 74)	40	25	4302	**3939**	363
419	(9, 8)	37	26	3858	**3665**	193
427	(32, 31)	5	58	**3664**	3677	-13
436	(24, 22)	64	4	**4050**	**4050**	0
439	(15, 14)	99	7	**4097**	**4097**	0
457	(6, 5)	46	8	**3774**	**3774**	0
490	(41, 40)	207	9	**4098**	**4098**	0
517	(22, 75)	66	24	**4090**	**4090**	0
520	(21, 51)	89	2	**4008**	**4008**	0
522	(39, 35)	67	7	4116	**3738**	378

Table 5. The best total costs calculated by RR1 and HMA after demand increased events in egl-e1-A

crc	*dit_arc*	*dit_dem_inc*	*dit_sc_inc*	RR1	HMA	diff.
326	(32, 34)	36	36	4318	**4037**	281
344	(54, 52)	3	3	**3551**	**3551**	0
345	(50, 52)	1	1	**3549**	**3549**	0
374	(52, 54)	2	2	**3550**	**3550**	0
376	(68, 66)	23	23	**3571**	**3571**	0
384	(44, 45)	6	6	3567	**3554**	13
415	(46, 47)	9	9	3975	**3557**	418
431	(44, 59)	11	11	3885	**3559**	326
449	(32, 35)	11	11	4280	**3559**	721
468	(35, 32)	65	65	4334	**4326**	8
490	(44, 46)	2	2	3636	**3550**	86
490	(32, 33)	28	28	4289	**3674**	615
493	(59, 44)	5	5	3879	**3553**	326
516	(35, 32)	24	24	4293	**3670**	623
545	(35, 41)	4	4	**3552**	3612	-60

Table 6. The best total costs calculated by RR1 and HMA after vehicle breakdown events in egl-e1-A

crc	*vb_id*	RR1	HMA	diff.
305	2	4124	**4076**	48
311	1	4060	**4012**	48
342	2	4204	**4122**	82
344	0	**4096**	**4096**	0
364	0	**4096**	**4096**	0
399	4	**4004**	4017	-13
451	2	4282	**4261**	21
463	0	3718	**3639**	79
490	2	4282	**4261**	21
495	3	**3548**	**3548**	0
506	0	3621	**3561**	60
507	1	3874	**3548**	326
523	2	**4261**	**4261**	0
540	0	**3548**	**3548**	0
541	2	**4282**	**4282**	0

CONCLUSION AND FUTURE WORKS

In this work, the possible events that change the current CARP instance were collected, then analyzed. Based on the results, a framework for data-driven DCARP was developed with three event-handling algorithms to handle the events that can make the original schedule infeasible or obsolete. To minimize the changes made in the route plans, a minimal rerouting algorithm, the RR1 algorithm was developed. RR1 aims to find the least cost route plans by utilizing the travel and service log and rerouting at most one route plan of the vehicles. The implemented algorithms were tested on a wide range of DCARP instances and the RR1 algorithm was compared with an existing solution. The results showed that RR1 can give solutions fast, that in around half of the cases have the same cost as the solutions given by the currently most accurate complete rerouting algorithm, the HMA.

In this work, RR1 was compared to only one complete rerouting algorithm. In the future, it will be compared with other algorithms as well. Furthermore, it will be tested with more realistic travel and service logs that do not follow the route plans completely.

Acknowledgements

Supported by the ÚNKP-20-3 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

The authors acknowledge financial support also from the Slovenian–Hungarian bilateral project “Optimization and fault forecasting in port logistics processes using artificial intelligence, process mining and operations research”, grant 2019-2.1.11-TÉT-2020-00113, and from the National Research, Development and Innovation Office–NKFIH under the grant SNN 129364.

References

[1] Maniezzo, V. (2004). Algorithms for large directed CARP instances: urban solid waste collection operational support. *UBLCS Techni-cal Report Series, Bolonha, Italy: University of Bolonha, 27*.

[2] Babaee Tirkolaee, E., Mahdavi, I., Seyyed Esfahani, M. M., & Weber, G. W. (2020). A hybrid augmented ant colony optimization for the multi-trip capacitated arc routing problem under fuzzy demands for urban solid waste management. *Waste Management & Research, 38*(2), pp. 156-172.

[3] Eglese, R. W. (1994). Routeing winter gritting vehicles. *Discrete applied mathematics, 48*(3), pp. 231-244.

[4] Fink, J., Loebl, M., & Pelikánová, P. (2021). Arc-routing for winter road maintenance. *Discrete Optimization, 41*, 100644.

[5] Golden, B. L., & Wong, R. T. (1981). Capacitated arc routing problems. *Networks, 11*(3), pp. 305-315.

[6] Lacomme, P., Prins, C., & Ramdane-Chérif, W. (2001, April). A genetic algorithm for the capacitated arc routing problem and its extensions. In *Workshops on applications of evolutionary computation* (pp. 473-483). Springer, Berlin, Heidelberg.

[7] Tang, K., Mei, Y., & Yao, X. (2009). Memetic algorithm with extended neighborhood search for capacitated arc routing problems. *Transactions on Evolutionary Computation, 13*(5), pp. 1151-1166.

[8] Consoli, P., & Yao, X. (2014, April). Diversity-driven selection of multiple crossover operators for the capacitated arc routing problem. In *European Conference on Evolutionary Computation in Combinatorial Optimization* (pp. 97-108). Springer, Berlin, Heidelberg.

[9] Chen, Y., Hao, J. K., & Glover, F. (2016). A hybrid metaheuristic approach for the capacitated arc routing problem. European Journal of *Operational Research, 253*(1), pp. 25-39.

[10] Ting, C. J., & Tsai, H. S. (2018). Ant Colony Optimization with Path Relinking for the Capacitated Arc Routing Problem. *Asian Transport Studies, 5*(2), pp. 362-377.

[11] Corberán, A., & Prins, C. (2010). Recent results on arc routing problems: An annotated bibliography. *Networks, 56*(1), pp. 50-69.

[12] Mourão, M. C., & Pinto, L. S. (2017). An updated annotated bibliography on arc routing problems. *Networks, 70*(3), pp. 144-194.

[13] Corberán, Á., Eglese, R., Hasle, G., Plana, I., & Sanchis, J. M. (2021). Arc routing problems: A review of the past, present, and future. *Networks, 77*(1), pp. 88-115.

[14] Wang, X., & Wasil, E. (2021). On the road to better routes: Five decades of published research on the vehicle routing problem. *Networks, 77*(1), pp. 66-87.

[15] Archetti, C., Guastaroba, G., & Speranza, M. G. (2013). Reoptimizing the rural postman problem. *Computers & operations research, 40*(5), pp. 1306-1313.

[16] Liu, M., Singh, H. K., & Ray, T. (2014). A memetic algorithm with a new split scheme for solving dynamic capacitated arc routing problems. In *2014 IEEE Congress on Evolutionary Computation (CEC)* (pp. 595-602). IEEE.

[17] Mei, Y., Tang, K., & Yao, X. (2013). Evolutionary computation for dynamic capacitated arc routing problem. In *Evolutionary Computation for Dynamic Optimization Problems* (pp. 377-401). Springer, Berlin, Heidelberg.

[18] Monroy‐Licht, M., Amaya, C. A., Langevin, A., & Rousseau, L. M. (2017). *The rescheduling arc routing problem*. International Transactions in Operational Research, 24(6), pp. 1325-1346.

[19] Padungwech, W. (2018). Heuristic algorithms for dynamic capacitated arc routing (Doctoral dissertation, Cardiff University).

[20] Padungwech, W., Thompson, J., & Lewis, R. (2020). Effects of update frequencies in a dynamic capacitated arc routing problem. *Networks*, *76*(4), pp. 522-538.

[21] Tagmouti, M., Gendreau, M., & Potvin, J. Y. (2011). A dynamic capacitated arc routing problem with time-dependent service costs. *Transportation Research Part C: Emerging Technologies*, *19*(1), pp. 20-28.

[22] Yazici, A., Kirlik, G., Parlaktuna, O., & Sipahioglu, A. (2013). A dynamic path planning approach for multirobot sensor-based coverage considering energy constraints. *IEEE transactions on cybernetics*, *44*(3), pp. 305-314.

[23] Tong, H., Minku, L. L., Menzel, S., Sendhoff, B., & Yao, X. (2021). A Novel Generalised Meta-Heuristic Framework for Dynamic Capacitated Arc Routing Problems. *arXiv preprint arXiv:2104.06585*.

[24] Dulai, T., Werner-Stark, A., & Hangos, K. M. (2018). Algorithm for directing cooperative vehicles of a vehicle routing problem for improving fault-tolerance. *Optimization and Engineering*, *19*(2), pp. 239-270.

Management, Economics, Business and Marketing (IAC-MEBM 2021)

Global Energy Crisis: Impact on The Global Economy

Peterson K.OZILI[a], Ercan OZEN[b]1

[a] Central Bank of Nigeria, Nigeria
[b] University of Uşak, School of Applied Science,Department of Banking and Finance, Uşak, Turkey, ercan.ozen@usak.edu.tr, eozentr@hotmail.com)

Abstract

This paper presents an overview of the 2021 global energy crisis. It has been shown that a combination of post-COVID recovery, depleted fossil fuel energy reserves and extreme weather conditions led to a surge in global demand for energy. Fossil fuel energy reserves were depleted and were not sufficiently replenished due to the need to transition from fossil fuel energy to renewables. Failure to replenish fossil fuel energy reserves led to unavoidable energy shortages. Energy demand rose to unprecedented levels and the shortage could not be filled by renewables, thereby triggering the energy crisis or energy supply crunch. The energy crisis or energy crunch witnessed in the UK and some parts of Asia and Europe showed how the energy transition can have ripple effects across every corner of society. The energy crisis demonstrates, in part, that renewable energy is not yet fully developed, and is not ready to meet the ever-growing global demand for energy by households and corporates. There is need to consider calls to slowdown the energy transition until the renewable energy sector is fully developed and ready to meet the ever-growing global demand for energy. Slowing down the transition to renewables is a step in the right direction in light of recent events.

Keywords: Energy Crisis, Energy Transition, Renewable Energy, Climate Change, Fossil Fuel, Net Zero, Emission, Carbon Neutrality, United Kingdom, China.

1. INTRODUCTION

Climate change and the need to decarbonize the energy ecosystem has led to the need to attain net-zero carbon emission. This agenda is supported by the United Nations' COP26 and the Paris Climate Accords. Fossil fuel energy has served mankind for more than 200 decades. Fossil fuel energy powered the industrial revolution of the 1700s to 1800s.

Recently, research has linked greenhouse gas emission and global warming to activities powered by fossil fuel energy (see Reijnders and Huijbregts, 2007; Karmaker et al, 2020; Martins et al, 2018). As a result, countries have taken steps to tackle climate change by reducing their stock of fossil fuel energy and increasing their stock of renewable energy resources (Lazarus and van Asselt, 2018; Leonard et al, 2020). This has led to the emergence of several innovations, policies and events aimed at growing the renewable sector while abandoning the fossil fuel energy sector.

Renewable energy has its merits. The most important merit is that renewable energy does not pollute the atmosphere unlike fossil fuel energy (Goldemberg, 2006). The transition to renewable energy has been widely discussed among policy makers, and investors are willing to invest in corporations that take into account ESG factors. Investment in the renewable sector is growing rapidly, and there have been considerable improvements in battery technology. Yet, battery technology is not yet sufficiently advanced to store large amounts of energy, and to provide the power needed to meet national energy demands or to serve the national grid. At the same time, the fossil fuel industry has been relatively ignored as investors are moving away from the sector, leading to a fall in fossil fuel energy reserves. This

* Corresponding author.

created a vacuum. Subsequently, the post-COVID recovery and the lifting of pandemic-era restrictions led to unprecedented increase in energy demand which could not be met by renewables that were either undeveloped or developing, and could neither be met by depleted fossil fuel reserves which had been ignored due to energy transition advocacy. Excess energy demand amid depleted global fossil fuel energy reserves led to an energy crisis in countries like the UK, China, Lebanon and many more.

The challenge that countries around the world now face is to cut fossil fuel emissions to ensure that net zero carbon emission targets are met while also ensuring that countries have guaranteed energy supply to avoid any unwanted socio-economic fallouts of energy shortages. The 2021 global energy crisis has shown that achieving such task is not very easy.

The rest of the paper is structured as follows. Section 2 discusses the spread of the energy crisis. Section 3 presents some country-specific cases. Section 4 highlights some general effect of the crisis on some sectors of the economy and section 6 concludes.

2. GLOBAL ENERGY CRISIS SPREADS

The price of Brent oil rose above $80 a barrel at the end of September 2021 for the first time in three years. This shows that global demand was rising faster than supply amid depleting energy inventories. The increase in oil price could add inflationary pressure to the global economy. The rising oil price was generating speculation that the oil industry isn't investing enough to increase supplies to reduce the high oil price. Due to the de-carbonization agenda, investors wanted high returns on investment instead of higher output. For this reason, they moved capital to renewable energy sectors that offered good returns.

2.1. Europe

In Europe, the price of electricity increased by over 200% in Spain, and similar increases were recorded across the EU, particularly in the winter season. Multiple factors were responsible for the increase in the price of electricity such as: low natural gas stockpiles, low overseas shipments, and sluggish maintenance work that put nuclear generators and other plants offline. Further increases in electricity prices sparked protests by consumers. To resolve the energy crisis in Europe, many European countries such as Spain, Italy, Greece and the UK, adopted national measures to avert the crisis, such as offering subsidies to energy providers and imposing price caps, in order to shield citizens from rising electricity costs as their economies recover fully from the COVID-19 pandemic.

2.2. Asia

In Asia, India and China witnessed unexpected shortage of coal for utilities. The Chinese government began to ration power supply to factories in multiple Chinese provinces since June. Meanwhile, in India, more than 70% of power is generated from burning coal. India faced coal shortages in August as the inventory of coal with power plants reached critically low levels. The total coal stock in India fell to 8.317million mt in August from 37.41 million mt in January 2021, according to India's Central Electricity Authority. The 8.317million mt can only meet energy demand for five days. The Indian government resolved the situation by diverting coal away from non-power uses so that coal supply is only available for power-uses. The implication is that non-power users of coal will not have access to coal energy. Indian households were also affected by the rise in the price of global oil products when demand for cooking gas, petrol and diesel grew to about 11% in August.[2]

[2] https://www.moneycontrol.com/news/business/explained-why-europe-and-china-are-facing-an-energy-crisis-and-how-it-affects-india-7523541.html

2.3. Africa

African countries are helping some European countries to build up their energy reserves amid the global shortage. For example, Spain relied on Algeria for some of its supplies of natural gas. Algeria supplies natural gas to Spain through Morocco while Morocco takes 7 per cent of the gas that is delivered to Spain through Moroccan pipelines since 1996. Recently, a diplomatic conflict between Algeria and Morocco led the two countries to sever ties with each other. This led Algeria to find an alternative route to supply natural gas to Spain, and this will come at higher cost to Spain thereby increasing the cost of natural gas in Spain. Moving towards Southern African, South Africa is the largest producer and exporter of coal in Africa. 86% of South Africa's electricity came from coal alone, compared to the global average of 34%, surpassing India which generates 71% of its electricity from coal. The G-20 has increased pressure on South Africa to reduce its coal resources so that the country will be COP26 compliant. Other African countries that are heavily reliant on coal for electricity generation include Botswana and Zimbabwe.

2.4. Middle East

In the Middle East, the energy crisis hit countries such as Lebanon. In Lebanon, the imported fuel reserves dried up. This made fuel become a scarce commodity in mid-2021. It led to power outage across the country, forcing people to adjust to life without electricity. Lebanon has capacity to produce around 3600 megawatts of electricity but it currently produces about 700 megawatts, which is less than 50%. The implication is that Lebanon will rely on imported energy from Jordan, Syria and neighbouring countries, and will be exposed to volatile global energy prices.

3. SPECIFIC COUNTRY CASES

3.1. China

The combination of a post-COVID recovery and unusually hot weather increased electricity consumption in China in mid-2021.China witnessed its worse power outage in a decade. Over 20 factories in China's 31 provinces witnessed unexpected episodes of shutdown and slowdown due to loss of power as the energy crisis affected businesses and households.[3] Millions of households in North-East China were without power for a few days while others were without power for many hours, making it difficult to find electricity to heat or light their homes.

Prior to the energy crisis, China consumed more than 3 billion tonnes of thermal coal annually, but the total coal inventory at major power-generation groups during the crisis, in September, was 11.31 million tonnes which is enough to meet demand for only two weeks.[4] Consequently, output, orders and employment fell in September 2021.Also, the gap between China's coal inventory and the daily coal consumption began to widen in April. Notably, the manufacturing purchasing manager's index (mPMI) in September fell below the 50-point mark to 49.6. This was the largest contraction in factory activity in China since the pandemic in February 2020. The fall in output was due to slowdown in high-energy consuming industries, such as factories that process metals and oil products. The Chinese manufacturing sector was also affected by rising costs, production bottlenecks and electricity rationing. To meet supply shortfall, China purchased coal from countries such as Mongolia, Indonesia and Russia to meet the surging domestic demand although it had banned coal import from Australia prior to the energy crisis. The ban of coal imports from Australia led to coal shortage of 35 million tonnes of high-quality coal in the year of the ban. Presently, 70 per cent of China's imported coal is from Indonesia with an energy value of 3,800 kcal/kg.

The energy crisis in China was caused partly by the government's plan to decarbonize Chinese heavy industries and reduce pollution in response to the climate emergency movement. The goal is to stop energy firms from producing electricity using dirty coal that pollutes the atmosphere so that China can achieve the carbon-neutral target by 2060.

[3] https://www.scmp.com/economy/china-economy/article/3150457/china-power-crisis-thermal-coal-inventory-nears-record-low
[4] https://www.scmp.com/economy/china-economy/article/3150457/china-power-crisis-thermal-coal-inventory-nears-record-low

The government imposed controls on emissions which increased energy production costs. The implication is that many energy producers will produce energy at a loss. It discouraged producers from increasing supply. Financial institutions also reduced lending to businesses that used coal to produce electricity. These series of actions led to energy supply bottlenecks, and highlights the economic consequence of the current de-carbonization and decoupling agenda on China's economy. The de-carbonization regulations in China caused the energy crisis in China because it reduced coal production. It is also important to note that energy crisis in China is not a production problem. Rather, the restrictions imposed on fossil fuel products by the government led to supply shortfalls which led to rising cost of electricity. Government restrictions on coal also made it difficult to replenish coal that have been used up.

3.2. UK

Prior to the crisis, the UK provided government policy support for the transition from the old economy (fossil fuel economy) to the new economy (renewables) in the UK. This led to capital flight from fossil fuel businesses and led to an influx of capital to the renewables energy sector. Investors began to invest in ESG compliant renewable energy companies in the UK. With insufficient capital for fossil fuel businesses, it became difficult for fossil fuel businesses in the UK to finance their operations and boost production. Some fossil fuel businesses sought bailouts and they didn't get it. It forced them to layoff many workers including truck drivers. Then the energy crisis became severe in August 2021 just after COVID-19 pandemic restrictions were fully lifted. It led to a surge in energy usage in the cold winter in the UK. The excess demand for gas could not be met by the depleted natural gas inventories and low wind speeds in the UK. This led to gas shortages. External factors such as Brexit also contributed indirectly to the UK energy crisis as it made it difficult for the UK to import cheap gas supply from neighboring European countries. Brexit also made it difficult for heavy-goods-vehicle (HGV) drivers to enter the UK to drive trucks and help with the wholesale distribution of gas to many gas stations. As a result, Brexit contributed to the shortage of HGV truck drivers in the UK at the time.

Presently, the UK does not have a homegrown renewable energy sector. The UK still relies on fossil fuels and has significant exposure to volatile global gas prices. Some European countries which the UK rely on for some of its gas supply witnessed severe gas production challenges or shortages. As a result, these countries had fewer stock of gas to export to the UK. There were also political issues such as the decision of the EU to block the Russia-linked Gazprom's Nordstream 2 pipeline from passing through Europe to sell oil to European markets. This further reduced Europe's stock of gas reserves, leaving little for export to other European countries like the UK. The combination of a post-COVID recovery and low wind speed during the cold winter increased the demand for gas in the UK. In August 2021, there was a surge in energy usage in the cold winter in the UK leading to a gas supply crunch. Due to the eminent supply shortages, wholesale gas prices rose by 250% in January, and by over 70 percent in August 2021. The UK regulator capped the price that could be charged to consumers by energy providers. This restriction then has led to the collapse of four small energy companies as they couldn't increase retail energy price to cover their cost of operation. In a nutshell, the energy crisis in the United Kingdom manifested as a gas supply crunch that was caused by a multiplicity of factors. To resolve the crisis, the UK government can either bail out the fossil fuel energy companies or remove the cap limit on gas prices to allow gas prices to increase even higher for customers. The government also announced that it will offer 5000 visas to HGV drivers so that they can help with the distribution of existing gas reserves to several gas stations.

CONCLUSION

This paper discussed the cause of the 2021 global energy crisis. The crisis was caused by a multiplicity of factors beginning with the de-carbonization campaign and the energy transition. The urgency to scale renewable energy led to a painful energy crisis today as most consumers still rely on fossil fuel to meet their energy needs. There are lessons to learn from the 2021 energy crisis. Subsequent climate action against fossil fuel will have disruptive effects on consumers during the transition process. Government support for clean energy via renewables will come at a cost during the transition, and governments must be prepared to mitigate the effect.

References

[1] Goldemberg, J. (2006). The promise of clean energy. Energy policy, 34(15), 2185-2190.
[2] Karmaker, A. K., Rahman, M. M., Hossain, M. A., & Ahmed, M. R. (2020). Exploration and corrective measures of greenhouse gas emission from fossil fuel power stations for Bangladesh. Journal of Cleaner Production, 244, 118645.
[3] Lazarus, M., & van Asselt, H. (2018). Fossil fuel supply and climate policy: exploring the road less taken.
[4] Leonard, M. D., Michaelides, E. E., & Michaelides, D. N. (2020). Energy storage needs for the substitution of fossil fuel power plants with renewables. Renewable Energy, 145, 951-962.
[5] Martins, F., Felgueiras, C., & Smitková, M. (2018). Fossil fuel energy consumption in European countries. Energy Procedia, 153, 107-111
[6] Reijnders, L., & Huijbregts, M. A. (2007). Life cycle greenhouse gas emissions, fossil fuel demand and solar energy conversion efficiency in European bioethanol production for automotive purposes. Journal of Cleaner Production, 15(18), 1806-1812.

Impact of Internal Audit on Financial Performance: Commercial Banks

Arjeta HALLUNOVI[a]1

[a] *Lecturer, Department of Finance-Accounting, Faculty of Business, University Aleksander Moisiu Durres*

Abstract

This paper studies the impact of internal audit on financial performance in the context of the banking sector in Albania. The impact of internal audit on financial performance is an interesting topic that has become the focus of many academic studies. Unlike other sectors that attract only the attention of investors and shareholders, internal audit in the banking industry is interesting to many parties, such as borrowers, depositors and policy makers. One of the reasons is the essential role of the banking industry in influencing the economic situation in the country. To investigate the effect of internal audit on financial performance where financial performance was taken as a dependent variable and internal audit was considered as an independent variable with indicators (audit independence, standards, control and competencies) the linear regression model was used (SPSS). The main finding of the paper shows that internal audit standards have a greater impact on financial performance than audit independence.

Keywords: Albania, Audit Independence, Audit Standards, Banking Sector, Internal Audit.

1. INTRODUCTION

The internal auditor plays the role of an important partner that helps the relevant organization meet its objectives by bringing a systematic and disciplined approach to evaluating and improving risk management, control and governance processes that lead to improved financial performance (Sauer, Schneider, Sheikh and Simione, 2013). The financial sector in Albania and more specifically commercial banks have experienced a contraction in the last two years 2017-2018. Their number from sixteen banks has gone up to twelve. The financial sector is seeing internal audit as an alternative way to influence the bank's financial performance. But, as in any endeavor, there are successes and failures, and the banking sector has demonstrated this quite well. The internal audit syndrome in the Albanian banking sector has entered in recent years. Although it is thought that internal audit leads to increased financial performance for the bank and their transformation into strong actors in the regions and the banking environment, the conditions in which this internal audit process is carried out affect and transform the banking sector and also affect confidence of customers, investors and shareholders, therefore should be valued (Bota and Palfi, 2009). The ability of banks to increase their market power and benefit from internal audit enhances their financial performance (Sikka, Filling and Liew, 2009). Financial performance requires appropriate risk-based auditing practices, ie an effective and efficient internal audit (Schneider, 2010). Internal audit also plays an important role in the stability, performance and ability of the bank to provide liquidity in difficult market conditions (Song, 2011). The purpose of this paper is to analyze the impact of internal audit on financial performance, specifically in a commercial bank in Albania, by how the implementation of internal audit affects financial performance.

2. LITERATURE REVIEW

In essence, auditing is represented as a process designed to assess the reliability and information of a company's financial statements (Carey, Subramaniam and Ching, 2006). Gacon (2013) argues that performing audit work according to international auditing standards contributes significantly to audit effectiveness. Carey, Subramaniam and

* Corresponding author.

Ching (2006) found that compliance with professional standards is the most important contributor to value added. According to Cai Arena and Azzone, (2009) independence is the essence of auditing. An international auditor should be independent of both the staff and the operational activities of an organization (Ima, 2015). Otherwise, the integrity of the auditor's opinions, conclusions and recommendations would be questionable (Zain and Subramaniam, 2007). Thus, independence is necessary for the effective achievement of the internal audit function and objective (Reinstein, Lander and Gavin, 1994). This independence is obtained mainly from two characteristics: organizational status and objectivity. Abbott, Parker and Peters (2012) in their attempt to provide evidence on the effectiveness of auditing in detecting errors affecting the financial statements suggest that all intentional errors are concentrated in a few audits and these are quite predictable by the industry. Most of these errors affect the revenues.
Adekola (2018) researched the impact of internal audit on the financial performance of commercial banks in Kenya. The study highlighted that financial performance requires a proper effective and efficient internal audit. From the findings, the study concluded that risk-based auditing through internal auditing standards and internal audit staff should be increased to enable firms to be able to detect risks in a timely manner and focus on areas of risk. high leading to increased transparency and accountability, increasing financial performance. This showed that there is indeed a relationship between internal audit and financial performance. Holm and Zaman (2012) emphasized that the existence of control is very important, especially in today's conditions with strong competition, which set a premium for reliable customer service, for the consideration of cash, for the realization of capital assets and manpower and for other reduction costs.

3. RESULTS OF THE ANALYSES

In this paper the time horizon used is cross-sectional. In this way, a particular phenomenon is observed at a given moment. In this case, the impact of internal audit on employee performance will be measured through questionnaires distributed to bank employees. The approach used in the paper is deductive positivist. The reason why this approach is the best fit is based on the arguments that what is required to be proven in this paper is based on existing theories that show that internal audit has an impact on financial performance and then finds results that support it. The questionnaire was distributed to 87 employees of the bank via email, of which 79 employees responded. Most of the participants belong to the range of 25-34 years, respectively with 56.3%. Regarding the level of education 74.1% of the participants have a master's degree. The work experience of the participants is in the interval 5-10 years with 47.1%.
Credibility - To test the reliability of the data, a Reliability Statistics Test was performed, which showed that the Cronbach Alpha coefficient has a value of 0.829. According to the test for the data to be reliable Cronbach's Alpha should take values above 0.7 to 0.9 and it is problematic when it takes values above 0.980. The coefficient in the case of data is at acceptable levels and it can be said that the data are reliable.

Table 1. Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Basedon Standardized Items	N of Items
.829	.863	4

Source: Author calculations

Validity of data - To measure validity a factor analysis test is performed, ie KMO and Bartlett "s Test. The paper seeks to measure the impact of financial performance through four constructs: competence, standards, control and independence of internal audit. Through the test it is noticed that KMO is 0.636, any value of KMO above 0.5 indicates that the data are valid and that it is being managed to measure what is intended to be measured.

Table 2. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.636
Bartlett's Test of Sphericity	Approx. Chi-Square	93.024
	Df	6
	Sig.	.000

Source: Author calculations

H_1: Independence of internal audit has a positive effect on improving financial performance.
To test the first hypothesis, a linear regression analysis was developed where as a dependent parameter is financial performance and an independent parameter is the independence of internal audit. The following table gives the value of R^2 which is 0.550. According to this result, 55% of the change in financial performance is explained by the independence of internal audit.

Table 3. Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.742[a]	.550	.535	.27845	.550	36.681	1	30	.000

Source: Author calculations

In the table below it is noticed that the confidence level is (sig = 0.000) and F is 36.573 this shows that the model is statistically significant.

Tabela 4. ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	2.745	1	2.745	36.573	.000[b]
	Residual	2.425	30	.078		
	Total	5.170	31			

a. Dependent Variable: Performance
Source: Author calculations

To summarize, the model prediction result is as follows:

$$Y = -1.175 + 1.142X \quad (1)$$

Which means that a one-unit increase in the level of internal audit independence will increase financial performance by 1.142 units. So, it can be said that the first hypothesis is proven.

Table 5. Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.175	.796		-1.477	.150
	Q_Pavaresia	1.142	.189	.742	6.056	.000

a. Dependent Variable: Performance
Source: Author calculations

H_2: Auditing standards have a greater impact than audit independence on financial performance.
To test the second hypothesis, a linear regression analysis was developed where as a dependent parameter is financial performance and an independent parameter is internal audit standards. The following table gives us the value of R^2. The value found is 0.635. According to this result, 63.5% of the change in financial performance is explained by internal audit standards. Recall that the independence of internal audit explained 55% of the change in financial performance.

Table 6. Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change

1	.797[a]	.635	.623	.25086	.635	52.160	1	30	.000

Source: Author calculations

The following table shows that the confidence level is (sig = 0.000) and F is 52.160 this indicates that the model is statistically significant. To summarize, the model forecast result is as follows:

$$Y = 0.891 + 0.705X \quad (2)$$

Table 7. Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.891	.382	.797	2.330	.027
	Q_Standartet	.705	.098		7.222	.000

a. Dependent Variable: Performance

Source: Author calculations

Which means that a one-unit increase in the level of internal audit standards will increase financial performance by 0.705 units.

CONCLUSIONS AND RECOMMENDATIONS

The impact of internal audit on financial performance is a topic which is recently gaining a lot of importance. Various researches put their focus on the impact that each of the four internal audit factors has separately on the financial performance where they differ from one industry to another or even within the same industry. Some studies show that the independence of internal audit has a positive impact on the financial performance of the organization and other studies show that some of the other factors mentioned have a negative impact on financial performance. As a result, organizations often find themselves in the dilemma of which of the following factors has the greatest impact, which should be given even more importance in order to increase financial performance. The above paper aimed to measure the impact that Internal Audit had on the financial performance of the banking sector. A structured measuring questionnaire was used, which was distributed to bank employees and the data collected from the questionnaires were analyzed with the statistical package SPSS version 20.

The findings of the paper showed that the independence of internal audit has a statistically significant positive impact on financial performance (sig = 0.000), R^2 = 55% which means that 55% of financial performance is explained by the independence of internal audit and that a a one-unit increase in the level of internal audit independence will increase financial performance by 1.122 units. From the testing of the second hypothesis it is noticed that 63.5% of the change in financial performance is explained by the internal audit standards. Recall that the independence of internal audit explained 55% of the change in financial performance. So, the second hypothesis is not accepted. Based on these findings, it can be said that Internal Audit has an impact on the financial performance of the banking sector in Albania. Future research can be done in a larger sample and to have more concrete results it is suggested that performance be measured not only based on employees' perceptions but also to analyze other indicators whether financial or they can be indicators based on estimates. of financial performance realized by the relevant institutions.

References

[1] Adekola, I. (2018). *Impact of internal audit function and efficiency on organizational performance*. Journal of Accounting Research.

[2] Abbott, L., Parker, S. and Peters, G. F. (2012). *Internal audit Assistance and External audit timeliness*. Auditing: A journal of practice and theory, 34(4):3-20.

[3] Arena, M. and Azzone, G. (2009). *Identifying Organizational Drivers of Internal Audit Effectiveness*. International Journal of Auditing. pp. 43–60.

[4] Bota, C. and Palfi, C. (2009). *Measuring and Assessment of Internal Auditor's Effectiveness.* Annals of the University of Oradea. Economic Science Series. pp. 784–790.

[5] Carey, P., Subramaniam, N., and Ching, K. (2006). *Internal audit outsourcing in Australia*. Accounting and Finance. pp. 11-30.

[6] Carey, P., Subramaniam,N., Ching, K.C.W.(2006). *Internal audit outsourcing in Australia*. Accounting and Finance, 46: 11-30.

[7] Gacon, T. (2013). *The professionalization of Internal auditing.* Journal of *Management and Business administration, Central Europe*, 21(4):55-68.

[8] Holm,C., and Zaman,M. (2012). *Regulating audit quality: restoring trust and legitimacy.* Accounting forum, 36:51-61.

[9] Ima, J. N. (2015). *Effective internal audit as a tool for fraud deterrence in Nigerian Organizations*. Humanity and Social Sciences Journal. pp. 73-80.

[10] Mihret D., and Admassu M. (2011). *Reliance of external auditors on internal audit work: A corporate governance perspective*. International Business Research. pp. 67-79.
[11] Sarens, G., and De Beelde (2006). *The relationship between internal audit and senior management: A qualitative analysis of expectations and perceptions*. International Journal of Auditing. pp. 219-214.
[12] Reinstein, A., Lander, G.H., and Gavin,T.A. (1994). *The external auditor's consideration of the internal audit function.* Managerial Auditing Journal, 9(7):29-36.
[13] Sauer, D.A., Schneider, G.P., Sheikh, A., and Simione, K.A. (2013). *Risk Management and the internal audit function in Sin firms.* Allied academies international conference, 18(1):65-68.
[14] Sikka, P., Filling, S.,and Liew, P. (2009). *The audit crunch: reforming auditing*. Managerial Auditing Journal, 24(2): 135-155.
[15] Song, C. (2011). *Problems and countermeasures on corporate internal audit in China*. Asian Social Science, 7(1):191-193.
[16] Schneider, A. (2010). *Assessment of Internal Auditing by Audit Committees*. Academy of Accounting and Financial Studies Journal. pp. 19–26.
[17] Vondra, A. and Schueler, D. (1993). *Can you Innovate your Audit*. Financial Executive. pp. 34–39.
[18] Zain M., and Subramaniam N. (2007). *Internal auditor perceptions on audit committee interactions: A qualitative study in Malaysian public corporations*. Corporate Governance 15, (5). pp. 894-908.

Challenges, Opportunities for The Development of New Banking Products in Pandemics

Egzona HASANI[a]1, Besim BEQAJ[b]

[a] *PhD. Candidate University of Pristina, Kosovo egzonahasani@hotmail.com*
[b] *University of Prishtina, Kosovo besim.beqaj@uni-pr.edu*

Abstract

The Covid - 19 pandemic has completely affected the economic system in the world and in Kosovo in particular. The anti-Covid measures taken also affected the functioning of the financial system and to a large extent commercial bank. At the same time it was an opportunity to create innovative products that meet the requirements of customers as a result of restrictions. The development of Smart products by banks was a big step characterized by competitive advantages and profit at the same time creating a stability sufficient operating and liquidity for banks. Research questions: How much has the pandemic affected the creation of competitive advantages of banks? Has the pandemic influenced the creation of smart products? The methodology consists in the collection of primary and secondary data. The tool for data collection is the structured questionnaire.

Keywords: New product development, financial system, covid 19, smart products, bank.

1. INTRODUCTION

Majority of business people, the ones who are decision makers, regularly think big finding best ways to strengthen their companies through updated, trendy and innovative products and services. Developing new product, although it's a hard process and takes great deal of risk, there are odds of being converted into huge benefits such is gain of market share, improving sales of existing products to retain customers or services that company must provide to increase funding and to keep its doors open. When companies consider that the best organization expect to close to half of all sales to come from new products commercialized in the last three years, than it is quickly to be realized that companies' developing next winning product or service can't be left to chance (Goldsmith, 2012). Where only few decades ago the focus was totally on the development of tangible products, nowadays it is about the development of ideas, plans, strategies and the creation of solutions instead. At the same time, there is growing emphasis on the contribution of products and services to the realization of the global sustainability. This means that new products should not be "new and improved", but "new to improve" the society in which they function (Joore, 2010). New products are central to the success of the modern company. For example, in 1986 it was estimated that 40% of manufacturers sales in the U.S. came from products that firms did not have five years previously. Further, a study of senior executives reveals that, by an eight-to-one majority, they believed their firms would be more depended on new products in the future. Coopers and Lybrand survey reported that most companies are counting heavily on new product development for growth and profitability in the years ahead (Cooper & Kleinschmidt**,** 2011). According to (James and Woelfel, 2000) new products represent the business opportunities for tomorrow, just as firm's existing products represent the business opportunities of today or the business opportunities of yesterday. According to some central banks, these negative processes transferred to the financial sector, mainly to banks, and resulted in a significant tightening of the credit policy, as well as a deterioration in the creditworthiness of borrowers, mainly from the SME sector.[2]

* Corresponding author.

[2] Alves, P.; Blanco, R.; Mayordomo, S.; Arrizabalaga, F.; Delgado, J.; Jiménez, G.; Asenjo, E.; Montes, C.; Trucharte, C. Recent Developments in Financing and Bank Lending to the Non-Financial Private Sector; Banco de

2. LITERATURE REVIEW

The pace of technology development and application has "globalized" communication, commercialization and has changed the industrial infrastructure through application software such are enterprise resource planning (ERP). It has created new industrial products and internet services that have replaced traditional banking services (e.g., ATM-s and online capability) (Rafinejad, 2007). New product development in a common context encompasses a large number of topics and challenges in a company; such are strategy formulation, deployment, resource allocation, and coordinated collaboration among people of different professions and nationalities, and systematic planning, monitoring and control. New product development has been an important topic for several business research disciplines, certainly economics, marketing, organizational theory, operations management and strategy (Loch & Kavadias, 2008). When company develops a new product or service, there is a process that is common to this development, and it can help ensure that you make the best product or service for your target audience. The challenge, however, is that some entrepreneurs or companies are tempted to skip one of these steps, in the hopes of making the process shorter, but that usually ends up short-changing the quality of the product or service they are developing. The reason is that product development costs rise greatly in later stages. Therefore, the company would like to go ahead only with those product ideas that will turn into profitable products (Claessens, 2015). Understanding the necessity of each phase of new product development can give the confidence to stick with it, even during the most frustrating moments (Quain, 2019). The strengthening impact of the bank's size, on the increase in credit risk costs caused by the COVID-19 pandemic, was also Sustainability 2021, 13, 11036 4 of 15 confirmed in the study by Borri and di Gorgio, conducted on a sample of listed European banks (Borri, N.; di Giorgio, G. , 2021). The banking sector concentrated disturbances experienced by both enterprises and households, including job cuts, a decline in commercial real estate prices, decrease in liquidity and corporate profitability, drop in central bank interest rates, and an increase in public debt. During the pandemic, the impact of these risk factors was particularly enhanced, and the likelihood of their materialization increased significantly. (Chetty, R.; Friedman, J.; Hendren, N.; Stepner, M., 2021).

3. STATISTICAL DATA AND RESULTS

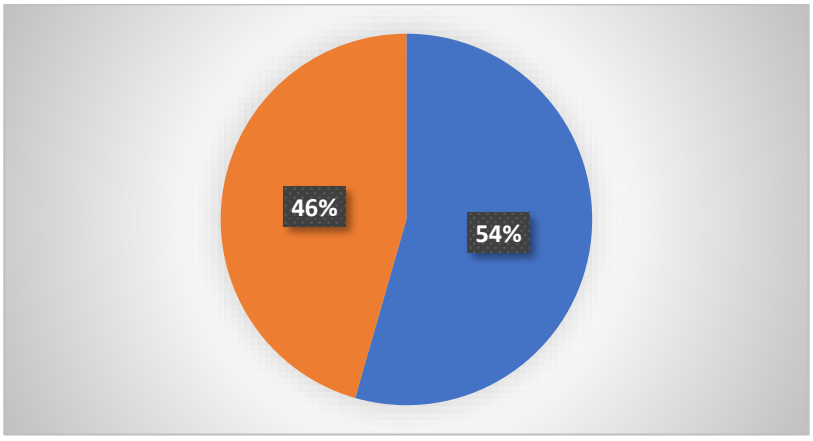

Fig 1. Areas of correspondents

In the figure of Areas is shown that there are 46% from rural areas and 54% from urban.

España: Madrid, Spain, 2021. Available online: https://repositorio.bde.es/bitstream/123456789/14751/1/be2101-art02e.pdf (accessed on 2 october 2021).

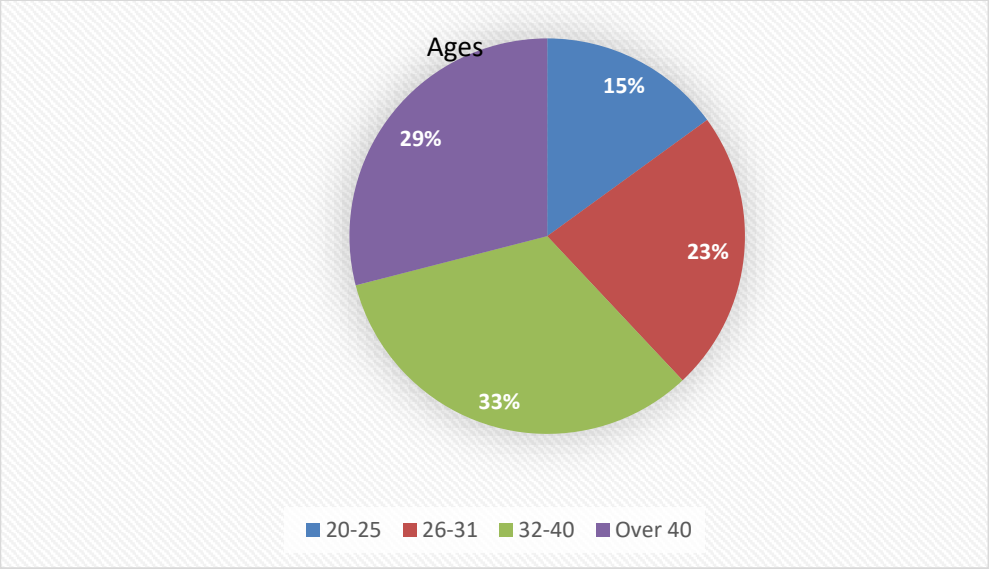

Fig 2. Ages for correspondents

In this figure are ages for correspondents: From 20-25 are 15%, from 26-31 are 23%, from 32-40 are 33% and over 40 are 29%.

Table 1. Definition coefficient of independent and dependent variables

Model Summary

Model	R	R Square	Adjusted R Square
1	.926[a]	.853	.838

a. Predictors: (Constant)? pandemic covid 19, bank strategy, information management system, innovation, plan proactivity

In this table is presented the coefficient of determination, R = 0.85, while Rsquare 0.84, or the adjusted coefficient 0.83X100 = 83%, which means the independent variables explain the dependent variable for 83%.

Table 2. significance expressed through anova the explanation F test

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	40.137	4	10.034	158.720	.000[b]
	Residual	18.650	295	.063		
	Total	58.787	299			

a. Dependent Variable: Smart products in Bank

b. pandemic covid 19, bank strategy, information management system, proactivity innovation

b. Predictors: (Constant),

Here is presented the model of Anova, where the level of significance is 0.000, which shows that the model has relevant links.

Table 3. Regression analysis expressing the impact of independent variables on dependent variable

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	.116	.066		2.755	.010
	pandemic covid 19,	.183	.074	.238	2.487	.013
	bank strategy	.590	.048	1.097	12.297	.000
	information management system	.862	.076	.483	11.410	.000
	Proactivity inovation	.179	.050	.373	3.607	.000

a. Dependent Variable: : Smart products in Bank

In this table is presented the regression analysis, which shows how many independent variables have an impact on the dependent variable, which is explained through the OSL equation,

$$\mathbf{Y=B_0+B_1X_1+B_2x_2+B_3X_3+B_4X_4+Ui}$$

In the table 1 are presented coefficients and their significance for the new banking products, the analyses done for banking market in Kosovo. Table shows that independent variables that consist: pandemic covid 19, bank strategy, information management system, proactivity innovation which all together have impact on dependent variable launching new banking products in pandemic times. Based in research question -Has the pandemic influenced the creation of Smart Products? conclusion derived are: Primary hypothesis is rejected and the second hypothesis is accepted, thus the interpretation of it provides a result. **Pandemic influenced the creation of SMART PRODUCTS.**

Table 4. correlation analysis expresses connectiveness between two variables.

Correlations

		Smart products in banks	competitive advantages in pandemics
Smart products in banks	Pearson Correlation	1	.432**
	Sig. (2-tailed)		.000
	N	300	300
competitive advantages in pandemics	Pearson Correlation	.432**	1
	Sig. (2-tailed)	.000	
	N	300	300

**. Correlation is significant at the 0.01 level (2-tailed).

According to this, there is a strong, positive and important relationship between smart products in bank and competitive advantages in pandemics at the level sig 0.000 which is less than 5% of the calculated coefficient. From this, we can say that there is a positive relationship between these two variables . **H2 Stands.**

CONCLUSIONS AND RECOMMENDATION

Banks in Kosovo need to create a cross-structural communication plan consisting of receiving up-to-date and accurate information about customer preferences and at the same time must develop modern innovative techniques in terms of launching contemporary smarts banking products. The development of global thinking approaches in pursuing strategies for launching new products in the market driven by technology, globalization, trade liberalization and homogeneity of customer demands and the pandemic situation.

- Banks in Kosovo should be oriented towards the development of the information system to create a plan in pandemic time.
- Buildings creative and innovative ideas to create a smarts products.
- Commercial banks should create a strategy which is compatible with customer preferences.
- The orientation of new products as competitive advantages in the market.
- Creating an efficient strategy which is oriented to digital marketing as a promotional instrument of the new products

References

[1] Alves, P.; Blanco, R.; Mayordomo, S.; Arrizabalaga, F.; Delgado, J.; Jiménez, G.; Asenjo, E.; Montes, C.; Trucharte, C. Recent Developments in Financing and Bank Lending to the Non-Financial Private Sector; Banco de España: Madrid, Spain, 2021. Available online: https://repositorio.bde.es/bitstream/123456789/14751/1/be2101-art02e.pdf (accessed on 2 october 2021).

[2] Cooper, G.R., & Kleinschmidt, J.E. (2011). "New products: the key factors in success." American Marketing Association, Chicago, Illinois 60606. USA.

[3] Claessens, M. (2015, June 24). New product development process (NPD) – obtain new products: Goldsmith, D. (2012). "*Creating new products and services.*" BenBella Books, Inc. Dallas, TX, 75321. USA.

[4] Joore, P. (2010). "New to Improve: The mutual influence between new products and societal change processes."VSSD, Leeghwaterstraat, 2628 CA Delft, The Netherlands.

[5] James, W. R., & Woelfel, J.C. (2000). "*The essentials of developing new products & services and operating profitably*." Research & Education Association, Piscataway, New Jersey 08854.US.

[6] Loch, C., & Kavadias, S.(Ed.). (2008). " *Handbook of new product development management(1st ed.).*" Butterworth Heinemann imprint of Elsevier Linacre House, Jordan Hill, Oxford, OX2 8DP, UK.

[7] Principles of marketing explained Marketing –insider. Retrieved from https://marketing-insider.eu/new-product-development-process/

[8] https://www.freiheit.org/western-balkans/impact-covid-19-economy-kosovo

ISBN 978-80-88203-24-7

Czech Institute of Academic Education z.s.