VOLUME 12 NUMBER 1 BANJA LUKA, JUNE 2022 (1-64) ISSN 2232-9625 (Print) ISSN 2233-0194 (Online) UDC 004

Journal of Information Technology and Applications

(BANJA LUKA)

JITA

Exchange of Information and Knowledge in Research

APEIRON

THE AIM AND SCOPE

The aim and scope of the Journal of Information Technology and Applications (JITA) is:

• to provide international dissemination of contributions in field of Information Technology,
• to promote exchange of information and knowledge in research work and
• to explore the new developments and inventions related to the use of Information Technology towards the structuring of an Information Society.

JITA provides a medium for exchanging research results and achievements accomplished by the scientific community from academia and industry.

By the decision of the Ministry of Education and Culture of the Republic of Srpska, no.: 07.030-053-160-4/10 from 3/3/2010, the journal „Journal of Information Technology and Applications“ Banja Luka is registered in the Registry of public organs under the number 591. Printed by Markos, Banja Luka in 300 copies two times a year.

Indexed in: LICENSE AGREEMENT, 3.22.12. **EBSCO** Publishing Inc., Current Abstracts

ebscohost.com	road.issn.org
citefactor.org/contact	citefactor.org
scholar.google.com	cosmosimpactfactor.com
doisrpska.nub.rs	
crossref.org	

Printed on acid-free paper

Annual subscription is 30 EUR
Full-text available free of charge at http://www.jita-au.com

UDC 004

ISSN 2232-9625 (Print)
ISSN 2233-0194 (Online)

JITA

VOLUME 12 NUMBER 1 BANJA LUKA, JUNE 2022 (1-64)

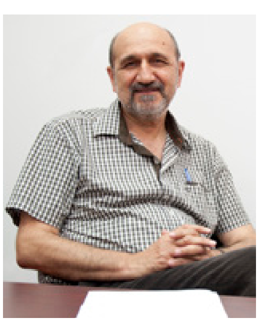
Contents

EDITORIAL

EDITORS:

ZORAN AVRAMOVIĆ, PhD
EDITOR-IN-CHIEF

GORDANA RADIĆ, PhD



DUŠAN STARČEVIĆ, PhD

Dear authors,
Dear readers,

Welcome to the twenty-third edition of the Journal of Information Technology and Applications (JITA), published by Pan-European University APEIRON Banja Luka.

The Journal of Information Technology and Application (JITA) publishes quality, original papers that contribute to the methodology of IT research as well as good examples of practical applications.

JITA is a scientific and professional information journal that aims to provide a forum for engineers, researchers and academics, universities and the industry to showcase their scientific research and encourage them to further learn and explore.

Our journal is independent of any external influences, does not charge any money from authors for publishing papers and does not operate on the principle of making a profit. Therefore, we are free to invite you to be a full participant in the creative process of spreading information knowledge.

Our vision, which we have been practicing for twelve years so far, is to create a quality publication that will be respected, appreciated, thought-provoking, which will encompass the different views of a wide range of authors.

We welcome the submission of the following in our JITA journal:

- original research,
- theoretical contributions,
- literature review,
- critical comments,
- case studies,
- book reviews, and
- current research.

We are glad that you have joined us as readers, and hope that you may join us as paper authors. We can proudly point out that JITA started with the selection of quality articles and has continued as such to this day. Members of the review team come from all continents, as well as the paper authors. With two independent "blind" reviews, we manage to select the highest quality articles and publish them afterwards.

Gratitude

On behalf of the Editorial Board, we would like to thank the authors for their high quality contributions, and also the reviewers for the effort and time invested into the preparation of the Journal of Information Technology and Applications.

Acknowledgments

We would like to thank many who have read and/or commented on earlier versions of this journal including Leonid A. Baranov, Yuri M. Inkov, Efim N. Rozenberg, Željko Stanković, Katarina Držajić Laketić and Dražen Marinković. However, any errors or shortcomings remain my full responsibility.

Conflicts of Interest

The author declares no conflict of interest.

Editor-in-Chief, Zoran Avramović
zoran.z.avramovic@apeiron-edu.eu
Editors, Gordana Radić, Dušan Starčević

Train command and control for commuter and urban lines

Efim Rozenberg[1], Alexey Ozerov[2]

Research and Design Institute for Information Technology, Signalling and Telecommunications on Railway Transport, 27, Bldg 1 Nizhegorodskaya str. Moscow 109029, Russia

[1]info@vniias.ru, [2]a.ozerov@vniias.ru

Contribution to the State of the Art

https://doi.org/10.7251/JIT2201005R UDC: 656.25.05:656.223.1.022.5

Abstract: The paper presents the state of the art of command and control and the challenges faced by the Russian Railways (RZD), with a focus on the migration to new paradigms of train separation, train localization and obstacle detection. The authors give an overview of the practical results of some ongoing projects carried out with the direct involvement of NIIAS researchers and developers for the Moscow Central Circle (MCC) railway.

Keywords: RZD; Virtual block; Moving block; Hybrid system; ATO; GOA3/4; Driver assistance system (DAS); Self-driving (autonomous) trains; Artifical Neural Network (ANN); GNSS; Digital route map.

Introduction

Today, the key task for RZD is to increase the capacity with simultaneous migration to adaptive traffic planning and management.

The importance of this task is due to the surging competition presented by other modes of transport, as well as by the substantial growth of agglomerations that creates an even greater load for railways up to the point of depletion of the design capacity.

Solving the problem most importantly requires further automation of the traffic management systems based on real-time supervision, conflict identification and resolution using artificial intelligence and advanced simulation tools. It also necessitates the development and application of new train spacing models enabling migration from conventional principles based on fixed blocks with trackside signals to virtual blocks with no trackside signals. Such models are especially applicable to high-density commuter or urban lines. The same holds for the migration to self-driving trains.

Implementing the above approach involves such prerequisites as the availability of reliable wireless communication infrastructure, high-precision train localization, advanced trackside train detection and on-board obstacle detection systems. Safety and security considerations are also of great importance [1].

Migration to new train spacing models

According to UN data, 55% of the world's population now lives in urban areas, and by 2050 the proportion is expected to rise to 70%. That means that two thirds of the humanity will live in urban agglomerations creating unprecedented pressure on the urban transportation networks, including the railways.

Every year, RZD carries over a billion passengers, and the largest share is carried by suburban lines in large urban agglomerations like Moscow, whose transportation infrastructure is experiencing great pressure.

To address the situation, in September 2016, a new type of urban railway system called the Moscow Central Circle (MCC) was put into revenue operation. The MCC is a 54-km-long above-ground orbital metropolitan passenger railway line that is partially integrated with the metro. Today, the MCC carries over half a million passengers per day. The figures are expected to double by 2030, though the six years of operation show that this will probably happen much earlier.

In many ways, the MCC serves as a site for testing and validating new technologies to be later deployed on mainlines of the network. For instance, at the MCC, the virtual block principle is implemented using short (300 – 400 m) audio frequency track circuits (AFTC) and GLONASS / GPS-enabled ATP units. Audio frequency track circuits enable train separation with no trackside signals and dynamic headway management.

Normally, each virtual block section is composed of 3 or 4 track circuits (TC). The boundaries of the so-called "moving" block sections are not strictly referenced to trackside physical assets like signals or board markers. They rather "follow" the tail of the train to ensure a safe distance between two consecutive trains. The codes of two adjacent track circuits may be different as they are generated according to the distance to the tail of the train ahead or to the entry signals at stations. Thus, the virtual boundary of a block section may fall within any pair of track circuits, and the resolution of train spacing is defined by the length of a track circuit rather than by the length of a block section made up of several track circuits [2].

As of today, the headway at the MCC has been reduced from the original 6 minutes further down to 4 minutes in peak hours using a high-precision coordinate system and a high-precision digital route map with so-called virtual balises, i.e. reference points in the map (see Fig. 1).

The boundaries of track circuits are used as reference points for virtual balises. The positioning errors are corrected at the moment of code changing and travelling over TC boundaries (the number and length of each TC is known and stored in the on-board database). Code changing in TC boundaries is identified by the on-board unit (OBU) with the precision of about 1 m. Reference points in the route map can be also put within a TC for further reduction of the distance between following trains down to the length of a braking distance plus safety overlap [3].

MULTI-LAYER COMMAND AND CONTROL

The above principle can also be implemented as an extended hybrid version of ERTMS Level 3 using Radio Block Centre (RBC) and radio communication over the national signalling system (as a fall-back). That would help reduce the headway down to 2 minutes.

That will represent further evolution for the MCC in the years to come, with the installation of a Russian version of RBC (or route server) communicating with the modified OBU through GSM-R or LTE. Now, radio is used for driver – dispatcher voice communication, as well as for delivering schedule corrections to the on-board driver assistance system, or the on-board automated train operation unit (ATO) while the safe route logic is executed by the automatic block (AB) system that manages the open lines, as well as main tracks within the MCC stations.

The automated traffic management system (TMS) supervises traffic at the MCC and automatically identifies and resolves conflicts by calculating and executing an alternative schedule. Train posi-

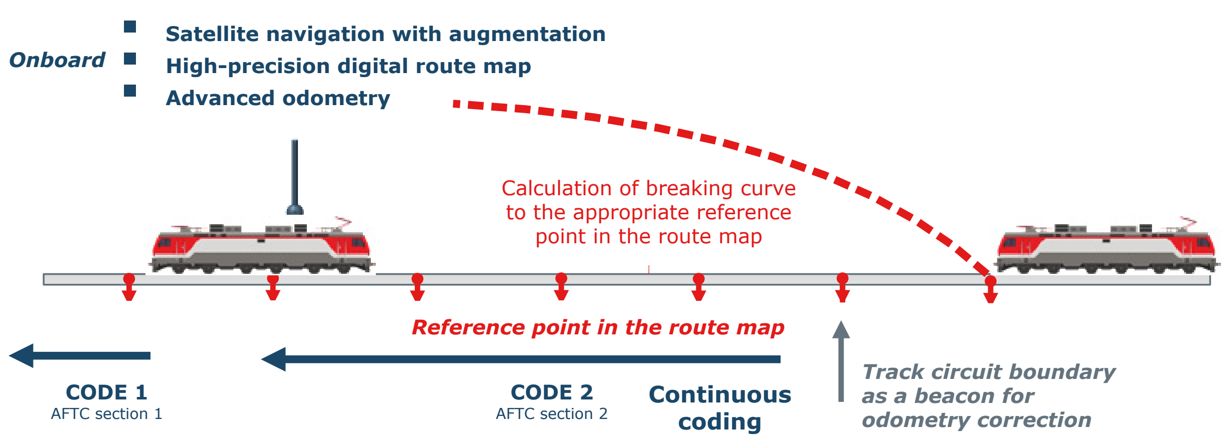


Fig. 1. *Virtual block sections at the MCC*

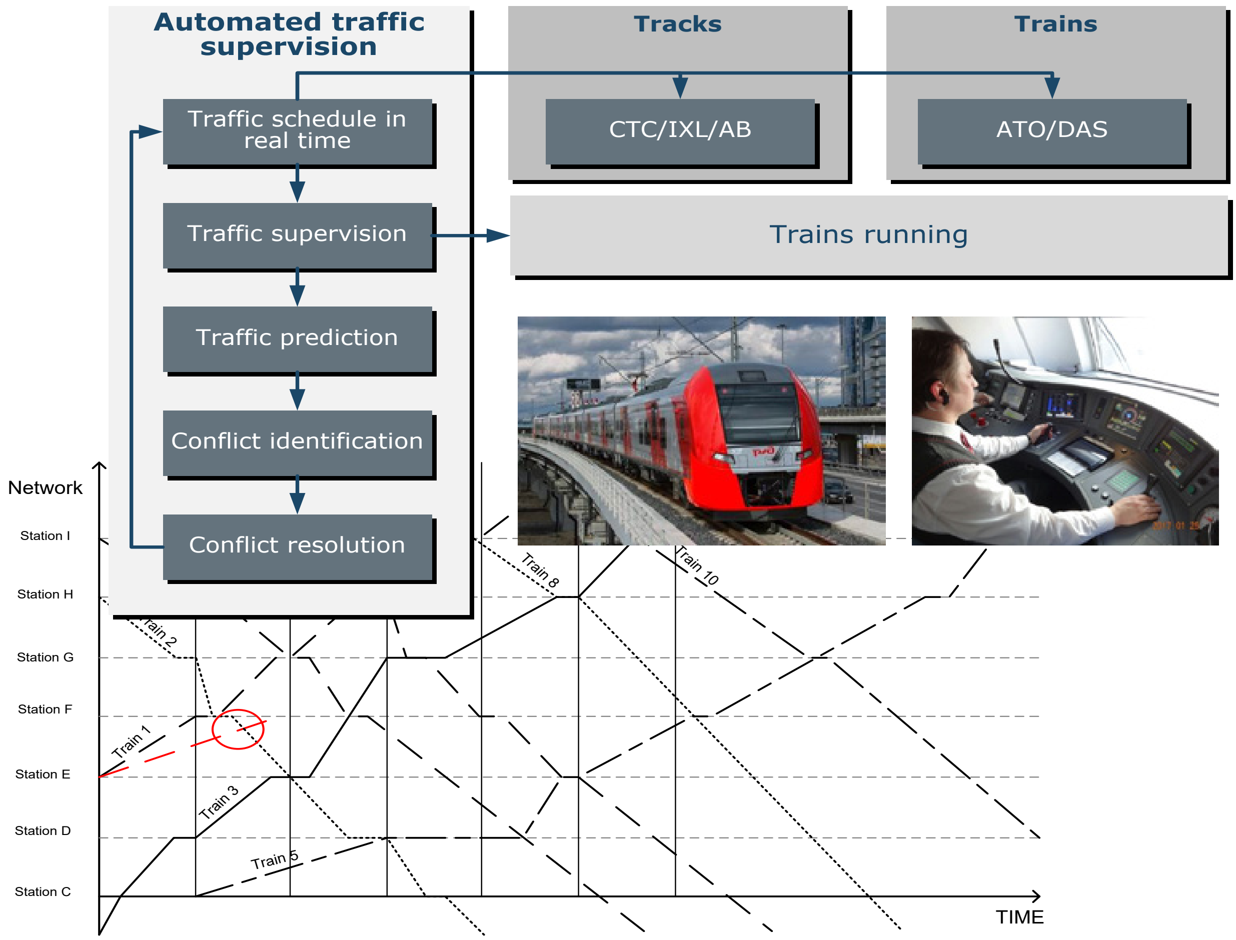


Fig. 2. Key TMS functionality

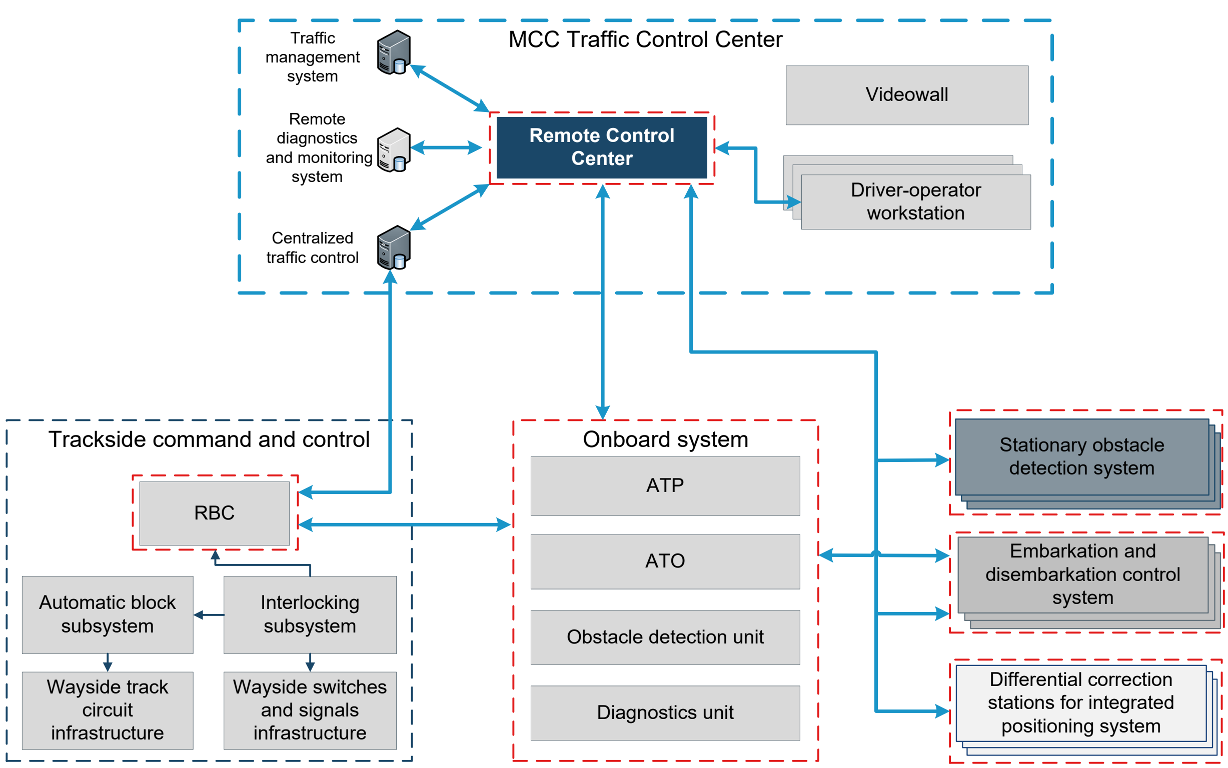


Fig. 3. Multi-layer command and control system with GoA3 / 4 functionality

tion and speed are supervised using GSM-R and an integrated localization system based on satellite navigation as part of the OBU. The TMS operation cycle is shown in Fig. 2.

Compared to subway systems where access to track is restricted and the boarding / alighting process is simplified by platform screen doors, urban railways have to resort to different solutions. Those include CCTV on platforms, trackside and on-board perception (automatic obstacle detection) subsystems used for protecting passengers on platforms, running trains, and people who may appear on the tracks.

Given the plan of putting self-driving EMUs in revenue operation at the MCC by the end of 2022, the general architecture of the MCC command and control is as follows (see Fig. 3):

Evidently, constructing safety models of such complex multi-loop transportation systems requires a comprehensive method. Such method must include systems analysis of unsafe scenarios along with the compilation of scenario library and formalization of a hazard model's description, pertaining to the boundaries of various control loops as well. The systems analysis may further result in the review and modification of the safety model of the transportation system under development and the conclusion regarding the requirement to include additional components into the model to perform the supervision and constraining function, e.g., by implementing a digital twin-based decision-making algorithm.

The introduction of perception (automatic obstacle detection) subsystems that use machine learning into the control loop significantly complicates the already challenging overall task of hazard analysis and safety evaluation of the multi-loop control system associated with the safety of people. The problem cannot be solved by means of the conventional FTA and FMEA hazard analysis methods only and requires a new comprehensive approach [4].

The MCC command and control system is designed as a multi-loop system that implies two control modes, i.e., autonomous and remote. In addition to the conventional track circuit-based train protection system, the control loop also supports radio communication between trackside and on-board train control and protection systems, as well as automatic obstacle detection by means of on-board and trackside perception modules that use ANNs and transmit relevant information to the remote control centre.

MIGRATION TO GOA3/4

Testing of automatic obstacle detection has been under way at the MCC for three years. The LTE network is used as part of remote train control to support the implementation of GoA3/4.

Currently, the Russian Railways is testing two types of trains equipped for GoA3/4. The trains are equipped with obstacle detection sensors and a unit for processing and communication with the remote control centre. The prototype trains are capable of automatic initiation of movement and braking in case of obstacle detection. However, during the testing, a driver is usually present in the cab and is always ready to take control. NIIAS's engineers also supervise the train's movement and monitor its systems in real time.

The first EMU was equipped in 2019 and is now used for data collection only. It is equipped with an infrared camera, LIDARs and eight optical cameras.

Based on the first test results, the second EMU was modified in late 2020 and is now used for testing a wide range of functions. The train is equipped with a new set of optical cameras, a LIDAR and infrared cameras with cleaning systems, short-distance ultrasonic sensors, an improved positioning system, pantograph and catenary monitoring cameras, as well as a boarding and alighting supervision system (Fig. 4).

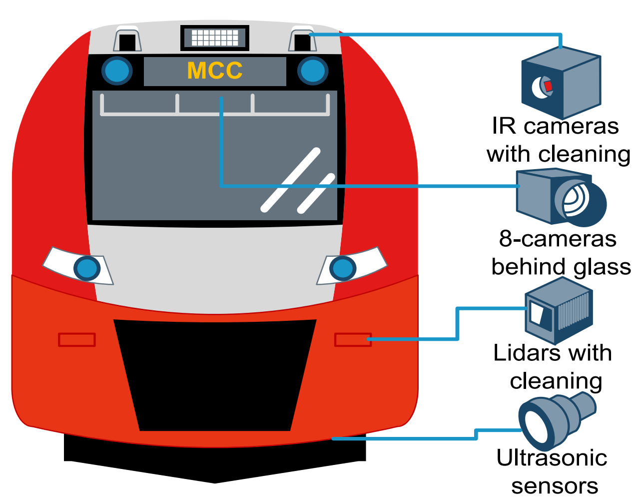


Figure 4. *MCC testing of autonomous EMU*

The trials primarily aim to test the four functions of the driverless system:

- detection of various classes of obstacles in varied weather and lighting conditions and on different infrastructure;
- remote train control;
- high-accuracy train positioning;
- video-based diagnostics of the pantograph and catenary.

Testing of the computer vision system's detection capabilities involves installing various obstacles ahead of the train, including human-like dummies. Up to 100 times a day, the modified EMU automatically detects dummies and stops before them in the morning, afternoon, evening and night lighting conditions, in a varied weather environment (Fig. 5).

The test results suggest promising opportunities of computer vision application. The calculated parameters show the system's superiority over a driver's eyes. For example, in daytime, the system detects a dummy at a distance of up to 600 m, while a person only sees it 400 to 500 m away. In night-time, both the system and the human driver recognize an obstacle from 250 to 300 m away; however, the on-board system has an advantage in the form of the infrared camera. In terms of the time of reaction to an obstacle, the average human performance is 1.3 sec, while the automated system's is 3 to 4 times better [5].

The multiple tests and the analysis of the sensors' efficiency in different conditions shows that there is no single perfect sensor, therefore the only possible solution is to combine data from various sensors using special algorithms, which allows eliminating the shortcomings of each individual sensor and combining their advantages, as well as supplementing them with information from other systems such as odometers, GNSS and on-board digital maps. The principal integrated solution is presented in Fig. 6.

Fig. 5. *Field testing of the prototype driverless train*

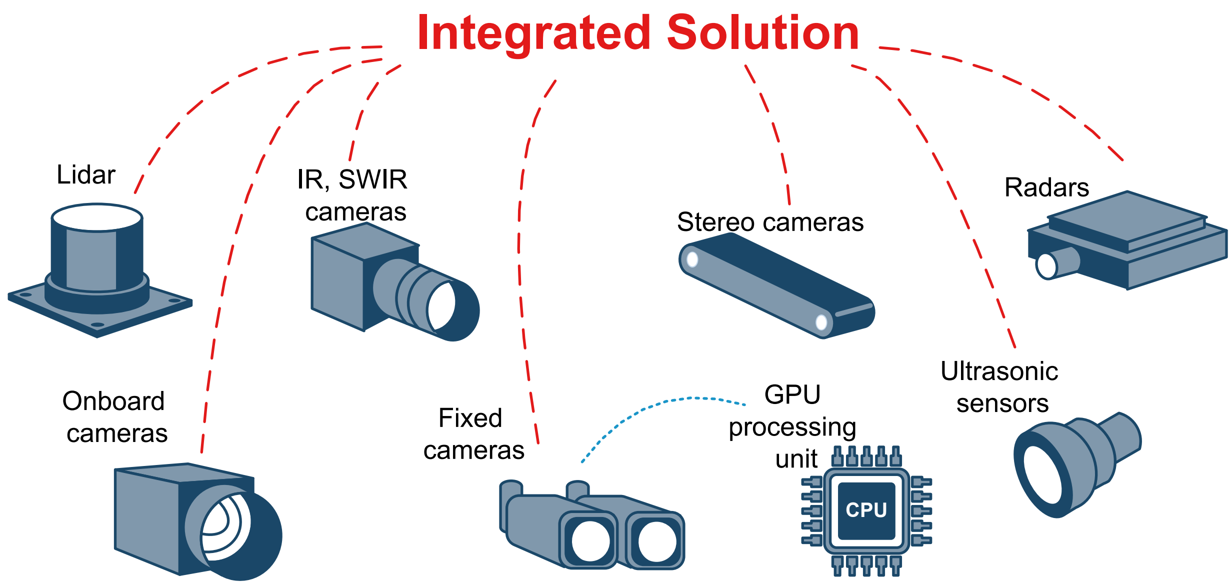


***Fig. 6.** Integrated solution with different sensor*

Detection of obstacles on the track is done by combining signals from several sensors. For each obstacle, a set of features is determined, such as:

- Coordinates of the object,
- Speed of the object,
- Dimensions of the object,
- Object class,
- Probability of object existence.

Objects are classified using neural networks and a signal is to be given if there are people close to the track (less than 5.5 meters from the axis of the track). The speed of objects allows predicting their future positions.

To implement automatic train operation, the obstacle detection system that receives data from sensors must also perform high-level tasks such as recognition, segmentation and identification. Algorithms that use artificial intelligence (AI) to solve these problems are based on machine learning (deep learning in particular), which requires massive labelled datasets.

The usual pipeline for AI data processing includes generation of data, filtering and labelling. The labelled data can then be downloaded into databases for further analysis or be used for different applications like test procedures, AI training or simulation (see Fig. 8) [6].

A typical machine vision dataset is designed for supervised machine learning. In addition to the input data received from the sensors, the dataset contains the target output data. The creation of the correct target output, the so-called annotation or labelling, is usually done by humans.

The typical objects that are labelled on the data for railway applications include:

- people,

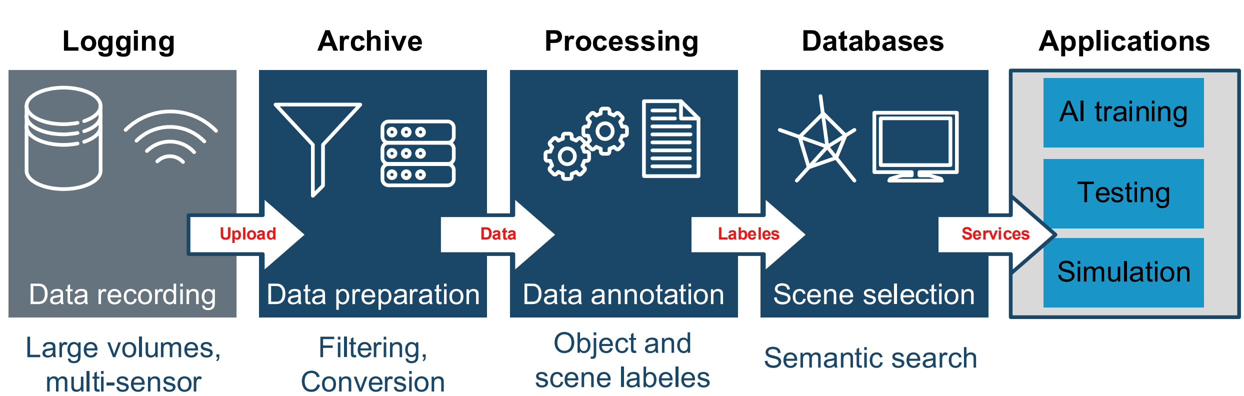


***Fig. 8.** Artificial intelligence data processing pipeline*

- railway workers,
- animals,
- vehicles / machines,
- supports (poles),
- signs and signals (with the status),
- switches.

Data labelling for computer vision can be of several types. The easiest option is to assign a category to the image. Another method involves manually placing image primitives, such as 2D bounding boxes, polygons or 3D regions (cuboids). With semantic annotations (or pixel-wise annotations) each pixel in an image is put in a category usually indicated by a unique colour code [7].

The data annotation process is usually rather costly and laborious since it is done manually. Despite the growing interest in AI-based applications in the last few years and the increased number of public datasets for road traffic applications, the railway industry has been lagging behind. The data used for the obstacle detection evaluation can be of three categories: publicly available data on the Internet, custom made datasets or data from real-world field trials. Relevant public datasets for railways are scarce and the majority of AI-based methods use custom-made and non-public datasets.

Another significant issue is that the initial data usually does not cover all possible cases and conditions. For that reason, augmentation, i.e., the creation of an artificial dataset based on the existing ones, is required. Augmentation mechanisms for railway-related images range from simple linear filters to complex models.

Another class of technical means required for successful operation of ATO systems are satellite positioning systems (GNSS) and platformless inertial navigation systems (PINS) that determine the current coordinates of the train and other railway infrastructure facilities with an accuracy of several metres, in combination with the use of a digital map. This significantly increases the efficiency of automatic control algorithms.

The practical use of an on-board digital map for autonomous driving is closely associated with the implementation of the map-matching algorithm. That is a method based on the image recognition theory that combines a digital map with information on the location of a train to obtain the actual position of rolling stock in the railway network. Map matching can be divided into two relatively independent processes: finding the path of movement of the current train and projecting the current location of the train onto such path.

INTEGRATED SOLUTION FOR TRAIN LOCALIZATION

It is worth mentioning that the Russian Railways was one of the world's first railway companies who started implementing satellite navigation in various railway applications as early as in the 1990s. Since then, all passenger trains and a considerable proportion of freight vehicles have been routinely equipped with GLONASS / GPS-enabled on-board ATP units, with the total number exceeding 20,000.

For the last ten years, all new passenger, freight locomotives and EMUs in Russia were equipped with the BLOK on-board system that combines the functionality of several subsystems and features the most innovative aspects of the OBU family while maintaining the highest safety integrity level (SIL4). BLOK has a modular architecture and can be customized for specific demands.

The design of the BLOK system allows using it both in stand-alone mode and jointly with other locomotive control and diagnostics systems. The system ensures train protection, including the cases when locomotives are operated by a single driver, without an assistant. The key functions of BLOK are as follows:

- receiving and processing of information from cab signalling systems, trackside devices and digital radio conduits (TLC unit),
- continuous speed monitoring and automatic brake application in case of overspeeding,
- prevention of SPADs,
- display of the target and permitted speed to the driver,
- rollaway protection,
- continuous driver vigilance monitoring,
- service and target braking,
- recording of vital train-borne information,
- display of movement parameters to the driver.

The simplified architecture of the BLOK unit is presented in Fig. 7 below.

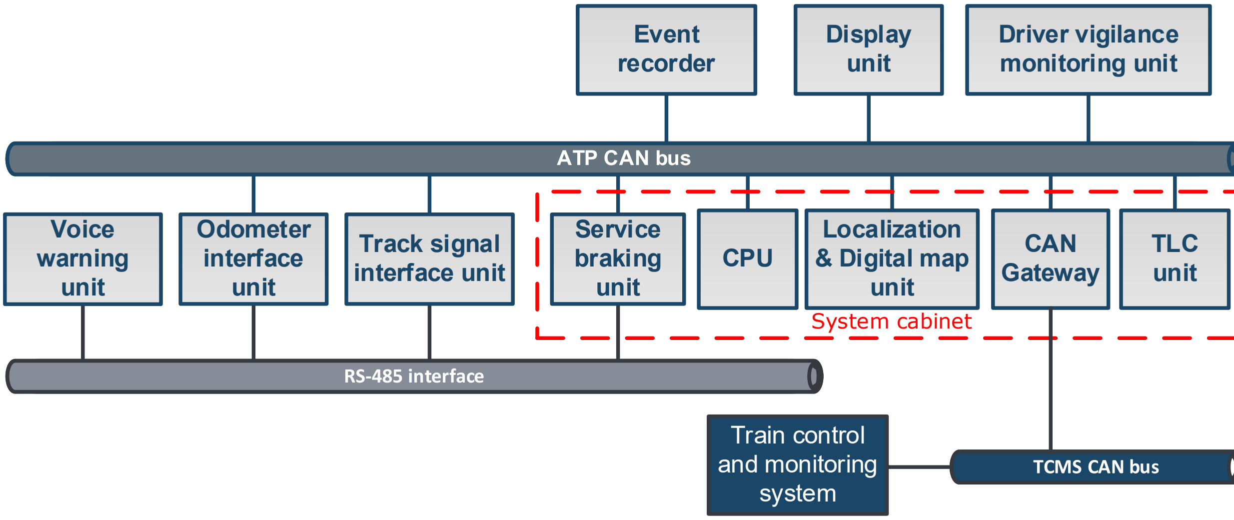


Fig. 7. BLOK simplified structure

A train localization unit in Russian ATP solutions typically uses GNSS positioning data in combination with an on-board digital route map. A standard train localization unit (TLU) with basic input data is presented in a simplified form in Fig. 8.

For the purpose of input data processing, the TLU uses a system of equations based on a Kalman filter to predict the system state evolution and to correct it accordingly. Combining GNSS positioning data with other independent data sources makes TLU outputs more reliable. Map matching helps reduce the search space to the actual route and translate the world coordinates into the linear ones to compute the distances to objects along the route stored in the map.

According to a number of sources [8], standalone GNSS-based positioning is not SIL4-safe. Obviously, it should be cross-checked with other reliable means like odometers, digital route map, track circuits or alternative facilities, and considered as part of some multi-layer railway system where not all of its constituents are to be necessarily SIL4. The safety integrity levels should rather be well balanced between them. In this context, a high-precision on-board digital route map plays an important role providing absolute positioning of authorized tracks, securing the continuity of the trip through previously validated localizations and improving the accuracy of speed profile and braking curve calculation.

A digital route map along with all the information it contains is included in the overall safety loop and acts as a reference, against which GNSS and other measurements are validated (provided that the accuracy of the digital map meets the operational requirements). As part of the navigation unit,

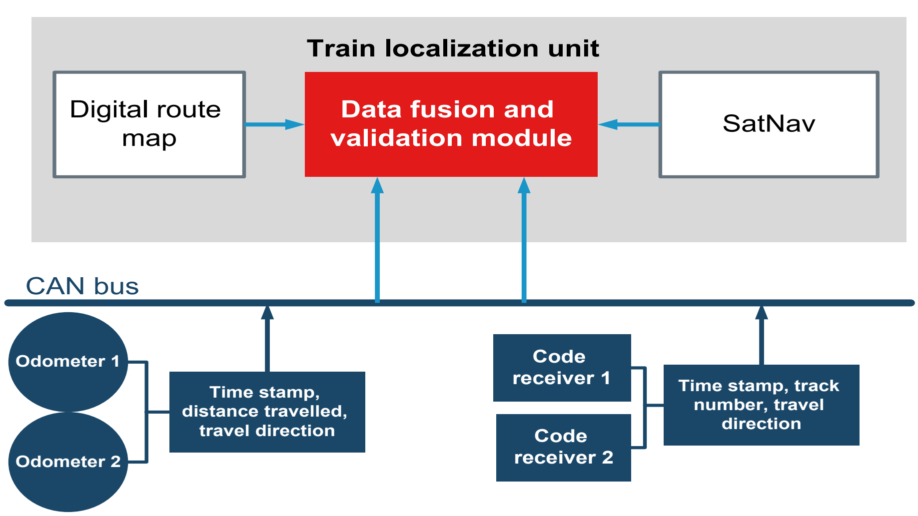


Fig. 8. Simplified structure of TLU

the digital map reduces human factor-related risks and serves as a source of information on trackside assets, the environment, the location of the vehicle even when there is no satellite navigation signal, the visibility is poor, etc.

Certainly, there is still much to be done to provide a more robust on-board train localization solution. Implementing a sound ATO functionality requires a long list of further improvements, including:

- automated generation, verification and updating of digital route map data,
- safe and secure methods of transmission / uploading of updated digital route database,
- standardized route map format and map matching mechanism,
- solutions for time synchronization of data received from different sources and integrated by the on-board data fusion and validation module,
- methods of in-motion verification of satellite navigation data,
- solution for start-of-mission positioning uncertainty,
- better satellite receivers with advanced in-built algorithms and shorter delay times.

Conclusion

The paper covered just a few ongoing developments enabling the migration to new models of train operation. On the face of it, it may seem that the examined issues are not closely related, but in fact, they are. The migration to GoA3/4 is not only about automatic detection of obstacles and replacing drivers with AI-based perception. It is also about absolute train localization and, therefore, safe operation. Ultimately, it is all about increasing capacity and punctuality.

References

[1] Rozenberg E., Ozerov A. Russia's digital railway begins to take shape. International Railway Journal 2017;57(4):38-40.

[2] Rozenberg E., Ozerov A. Digitisation of train command and control. In: Transportation Research Procedia 2019;40:1533-1539.

[3] Rozenberg E., Ozerov A. Contrôle-commande des trains dans les chemins de fer russes. RGCF 2021;320:31-41.

[4] Ozerov A., Olshansky A. Safety model construction for a complex automatic transportation system. Dependability 2021;2:31-37. https://doi.org/10.21683/1729-2646-2021-21-1-31-37.

[5] Popov P., Okhotnikov A. Autonomous train – the Russian perspective. Automation, Communication and Informatics 2019;8. (in Russ.).

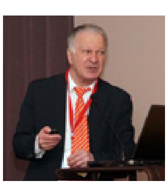
[6] Nieto M., Senderos O., Otaegui O. Boosting AI applications: Labeling format for complex datasets. SoftwareX 2021;13.

[7] Tagiew R., Buder T., Tilly R., Hofmann K., Klotz C. [Data set for autonomous driving as basis for GoA3+]. Railway technical review 2021;9. (in Germ.)

[8] A Survey of Train Positioning Solutions. IEEE Sensors Journal 2017;20.

Received: January 7, 2022.
Accepted: May 9, 2022.

About the authors



Efim Rozenberg is Doctor of Engineering, Professor who has a long-standing experience in signalling with about 40 years in railway research and development and at various managerial positions in railway R&D institutions of the Russian Railways. He is author of 300 patents and over 200 papers and one of the pioneer researchers in the field of GNSS application for train localization. He is honored member of the Russian Academy of Electrotechnical Sciences, member of the Russian Transport Academy. Currently, he is First Deputy of Director General of JSC NIIAS, R&D Institute for Information Technology, Signalling and Telecommunications on Railway Transport.

Alexey Ozerov is the Head of International Department of JSC NIIAS. He has been working with JSC NIIAS for over 15 years in various positions related to research, signalling business unit and international cooperation. He is finishing his PhD in traffic management and has authored about 30 papers and 7 patents in the field of advanced railway signalling and traffic management. He is JSC NIIAS representative in UIC, member of UIC Rail System and Cybersecurity Platforms, expert of IEC/TC 9.

For citation

Efim Rozenberg, Alexey Ozerov, Train command and control for commuter and urban lines, *JITA – Journal of Information Technology and Applications, Banja Luka*, Pan-Europien University APEIRON, Banja Luka, Republika Srpska, Bosna i Hercegovina, JITA 12(2022) 1:5-13, (UDC: 656.25.05:656.223.1.022.5), (DOI: 10.7251/JIT2201005R), Volume 12, Number 1, Banja Luka, June (1-64), ISSN 2232-9625 (print), ISSN 2233-0194 (online), UDC 004

SEMANTIC NUMERATION SYSTEMS AS INFORMATION TOOLS FOR FUZZY DATA PROCESSING

Alexander Yu. Chunikhin

Palladin Institute of Biochemistry, National Academy of Sciences of Ukraine, alexchunikhin61@gmail.com; ORCID 0000-0001-8935-0338

Vadym Zhytniuk

Palladin Institute of Biochemistry, National Academy of Sciences of Ukraine, v.zhytniuk@kau.edu.ua; ORCID 0000-0003-4034-3393

Contribution to the State of the Art

https://doi.org/10.7251/JIT2201014C UDC: 004.6.056.55:519.852

Abstract: We describe the concept of semantic numeration systems (SNS) as a certain class of context-based numeration methods. The main attention is paid to the key elements of semantic numeration systems - cardinal semantic operators. A classification of semantic numeration systems is given. The concept of fuzzy cardinal semantic transformation as a basis for creating fuzzy semantic numeration systems is advanced. Both fuzziness of the initial data - cardinals of abstract entities - and fuzziness of the parameters of the cardinal semantic operators are considered. The principle of formation of the fuzzy common carry in the cardinal semantic operators with multiple inputs is formulated.

Keywords: Cardinal Abstract Entity, Cardinal Semantic Operator, Semantic Numeration System, Fuzzy Cardinal Semantic Transformation.

INTRODUCTION

Any data processing in modern information systems is ultimately based on the use of one or another numeration system. A numeration system is a symbolic method of representing numbers using signs. Modern numeration systems are usually divided into three classes: positional (Place-value), non-positional and mixed. Despite a significant variety of works in the field of positional numeration systems [1-3], we can say that the overwhelming majority of them do not go beyond the traditionally linear representation of numbers.

“Place-value” is a representation of an abstract number by a system of characters (digits) - the contents of places - which are named in a certain way (by numbers, symbols or words). The semantics of the traditional place-value representation can be expressed as follows: *n* units of some abstract entity *i* are given the meaning (~>) of a unit of another abstract entity *j*: $n \cdot 1_i = n_i \sim> 1_j$, and further, recursively: $m \cdot 1_j = m_j \sim> 1_k$. Such representation, in which a certain amount of *one* entity is associated with a unit of *another* entity, can be called a linear or (1-1)-representation.

Assume that there are such abstract entities *i*, whose *n* units n_i are given the meaning of both the unit of an abstract entity *j* (1_j) and the unit of an abstract entity *k* (1_k) simultaneously: $n_i \sim> (1_j, 1_k)$. Consider another situation: to form a unit of an abstract entity *k*, exactly *n* units of an abstract entity *i* and *m* units of an abstract entity *j* are required: $(n_i, m_j) \sim> 1_k$.

Aren’t linear numeration systems the methods of constraining efficient data/information processing by reducing the semantic content of the data to its numerical value?

The author's previous works [4, 5] were an attempt to construct semantic numeration systems without cardinal semantic operators, but only on the basis of cardinal abstract entities connected in some topology. Abstract entities were given an active role both in the formation of the carry and in the formation of a structure of connectivity (entanglement) with other abstract entities.

PRELIMINARIES [6, 7]

Understanding *the entity* as something distinguished in being and having some meaning, we will introduce several definitions that concretize the application of this concept in the area covered.

An abstract entity (Æ) is an entity of arbitrary nature, provided with an identifier name that allows it to be distinguished from other entities. For example, a number, a car, a galaxy. Name as entity identifier can be either elementary (i) - one-element (letter, digit, word, symbol), or complex (composite, multi-element), corresponding to abstract coordinates (<i |) of the entity $Æ_{<i\,|}$ in some semantic variety. For example, "the zero bit of a number in binary notation."

Manifold (Multeity) is the next concept that is important for further presentation. From the many different definitions of this concept, we synthesize the following. *Multeity* - a manifestation of something uniform in essence in various kinds and forms, as well as a quality or condition of being multiple or consisting of many parts.

Since in what follows we will deal with the transformation of meanings, we will define the corresponding specific type of multeity as semantic. *Semantic multeity* is an abstract space with no more than a countable set of abstract entities, semantically united by the unity of the goal description (context). A semantic multeity will be called *open* if the number of abstract entities in it is countable (at least potentially), *closed* if it contains a finite number of abstract entities and *bounded* if the number of abstract entities in it is finite and unchanged.

Differences between semantic multeities and manifolds in mathematics are:

- semantic *heterogeneity* of entities that make up semantic diversity;

- *goal-setting*: a semantic multeity initially includes only those abstract entities that are used (potentially can be used) to solve a specific problem;

- *openness*: the possibility of new abstract entities generation as a result of transformations. In the end, a semantic multeity can form a semantic universe that includes any conceivable abstract entity.

Cardinal Semantic Multeity (CSM) is a semantic multeity, each element of which is equipped with a cardinal characteristic - the multiplicity of a given abstract entity is represented in multeity. From a set-theoretic point of view, a cardinal semantic multeity is a multiset, the carrier of which is contextually conditioned. The elements of cardinal semantic multeity will be called *cardinal abstract entities*.

Cardinal Abstract Entity (CÆ) is an abstract entity with a cardinal characteristic $CÆ_i = (i; N_i)$, where *i* is the name of the cardinal abstract entity, $N_i = \mathrm{Card}\,(CÆ_i) = \#\,(1_i, \dots, 1_i)$, $N_i \in \mathbf{N}$. We will assume that the named unit 1_i is a quantum of the meaning for the abstract entity $Æ_i$.

THE CONCEPT OF A CARDINAL SEMANTIC OPERATOR

This paper can serve as a certain theoretical supplement advantage to the work [8], in which the concept of an operator is given the status of conceptual in theory and practice.

It seems that there is no strict uniform definition of a *semantic operator* yet. Each application area - logic, linguistics, and programming - interprets this concept in its own way. However, it is possible to single out a certain semantic invariant that allows you to define the action of the semantic operator (SO) as a change/transformation of the meaning of a certain entity ($Æ_i$) or a set of entities ($Æ_i, \dots, Æ_k$) into another meaning ($Æ_j$): $Æ_j = SO(Æ_i)$ or

$Æ_j = SO(Æ_i, ..., Æ_k)$.

Let us introduce the following basic concept - the concept of *a cardinal semantic operator*. In essence, the action of the cardinal semantic operator is to give a certain number of units n_i of the cardinal abstract entity $CÆ_i$ the meaning of unit 1_j of the cardinal abstract entity $CÆ_j$, $(i \neq j)$: $n_i \sim> 1_j$. In principle, other options are also possible, for example, when the n_i of an abstract entity $CÆ_i$ is assigned not one, but simultaneously several different semantic units of respectively different CÆs: $n_i \sim> (1_j,..., 1_k)$. Or, for the generation of the unit of meaning 1_j of the abstract entity $CÆ_j$, the corresponding *n*-s of several other CÆs are simultaneously needed: $(n_i,..., n_k) \sim> 1_j$.

A *Cardinal Semantic Operator* is a multivalued mapping of the cardinal semantic multeity on itself, which associates a set of entity operands from the multeity with a set of entity images from the same multeity, transforming their cardinals using the operations defined by the operator signature: $Signt(CSO) = (K, Form, |n>_W, |r>_V)$, where K is a kind of operator, Form is a type of operator, |n> is a radix vector, and |r> is a conversion vector. The pair (W, V) is a valence of the Cardinal Semantic Operator.

Depending on whether the cardinal value of CÆ-operands changes under the action of the cardinal semantic operator, the latter can be representatives of one of three families - *transforming* operators (.), which change the value of the cardinals of both CÆ-operands and CÆ-images; *preserving* operators [.], changing the value of cardinals of CÆ-images and not affecting the value of cardinals of CÆ-operands; and *complex* operators (.], acting on some CÆ-operands as transforming, and on others - as preserving. This approach determines the possibility of the existence of three classes of semantic numeration systems - transforming, preserving and complex.

The kind of the cardinal semantic operator indicates the content of the transformations (definition of carry and remainder), which are performed both with the cardinals of CÆ-operands and with the cardinals of CÆ-images.

For example, the following kinds of CSOs are possible:

- radix-multiplicity (**↑#**): carry $p_i = \lfloor N_i/n_i \rfloor$, remainder $rem\ N_i = N_i - p_i n_i$;
- radix excess value (**↑Δ**): $p_i = N_i - n_i$; $rem\ N_i = 0$;
- radix excess fact (**↑•**): $p_i = 1 \Leftrightarrow N_i > n_i$; $rem\ N_i = f(N_i, n_i)$;
- arbitrary function (**f**): $p_i = f(N_i, n_i)$, $rem\ N_i = g(N_i, p_i, n_i)$.

In this paper, we consider the transforming cardinal semantic operators of the radix-multiplicity kind.

We will call the number of CÆ-operands of a cardinal semantic operator its *input valence* (W), $W = \#(CÆ_i,..., CÆ_j) = dim(|n>)$, and the number of CÆ-images - its *output valence* (V), $V = \#(CÆ_k,..., CÆ_l) = dim(|q>)$.

Strictly speaking, the output valence of the transforming CSO is determined by the sum of the actual output valence and the valence of "feedbacks", returning remainders to CÆ-operands. Then the full valence (–arity) of the transforming CSO will be (W, V + W). However, doubling the output valence in practice can lead to confusion, so we will usually neglect the valence of the return of reminders and write the (full) valence of the transforming CSO as (W, V).

Radix-vector $|n>_W = (n_i,..., n_j)^T$ consists of the particular radices (bases) $n_i,..., n_j$, relative to which the particular i,..., j-carries $p_i,..., p_j$ are formed. They further participate in the formation of the *common carry p* (if it's necessary).

The conversion vector $|r>_V = (r_{\cdot k},..., r_{\cdot l})^T$ consists of the components that determine the "scale factors" of the transformation of the common carry into the components of the *transformant* $q_k,..., q_l$, which change the values of the cardinals of the CÆ-images. This means that, for example, the carry *p* formed according to a given rule will be associated not with 1_j of the cardinal abstract entity $CÆ_j$, but with $r_{.j}$ of such units. We will call $r_{.j}$ the *rate of conversion* (j-conversion) of the carry. The introduction of conversion rates allows you to create numeration systems with rational bases.

The transformant $|q>_v = (q_k,\dots, q_l)^T$ is a direct result of the action of the CSO on the CÆ-operands. However, we will consider the transformation complete only after the "recalculation" of the cardinals of all CÆ-images of the operator in accordance with the obtained values of the corresponding components of the transformant $|q>_v$.

Specific values of valency, components of the radix vector $|n>_w$ and conversion vector $|r>_v$ are determined by the *specification* of the cardinal semantic operator.

The case when the sets of CÆ-operands and CÆ-images of an operator are singleton corresponds to the generally accepted positional numeration systems.

Forms of cardinal semantic operators

Let us define the *main forms* of cardinal semantic operators of the radix-multiplicity kind.

L-operator (Line-operator): $(\uparrow\#, L, n_i, r_{ij})$ – a cardinal semantic operator of valency (W, V) = (1, 1), which assigns (gives the meaning of) r_{ij} units of the transformant q_j, added to the cardinal N_j of the abstract entity $CÆ_j$, to each n_i of the cardinal abstract entity $CÆ_i$.

A schematic representation of the L-operator is shown in Fig.1.

Figure 1.

When an L-operator acts on a $CÆ_i$-operand, the following operations are performed:

(i) $p_i = \lfloor N_i/n_i \rfloor$ – calculation of radix-multiplicity, that is, i-carry value;

(ii) Rem: $N_i` = \text{rem } N_i = N_i - p_i n_i = N_i \bmod n_i$ – finding the remainder in $CÆ_i$;

(iii) $q_j = p_i \cdot r_{ij}$ – calculation of the j-transformant value;

(iv) $N_j` = N_j + q_j$ – finding the change of the $CÆ_j$-image cardinal.

The L-operator of signature $(\uparrow\#, L, n_i, 1_j)$ is basic for many traditional positional numeration systems. In particular, for the decimal numeration system: $(\uparrow\#, L, 10, 1)$.

D-operator (Distribution operator): $(\uparrow\#, D, n_i, (r_{ij}, \dots, r_{ik}))$ – a cardinal semantic operator of valency (1, v), which assigns the following units to each n_i of the cardinal abstract entity CÆi *v* transformants: r_{ij} units of j-transformants q_j for a cardinal abstract entity CÆj,…, and r_{ik} units of k-transformants q_k for the cardinal abstract entity $CÆ_k$.

A schematic representation of the D-operators D_2 is shown in Fig.2.

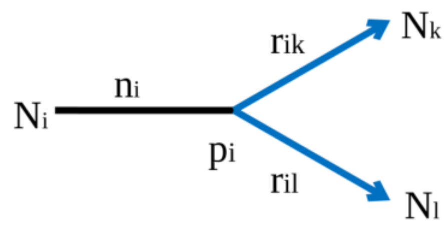


Figure 2.

When a D-operator acts on a $CÆ_i$-operand, the following operations are performed:

(i) $p_i = \lfloor N_i/n_i \rfloor$ – determining radix-multiplicity (i-carry value);

(ii) $N_i` = \text{rem } N_i = N_i - p_i\, n_i = N_i \bmod n_i$ – finding the remainder in $CÆ_i$;

(iii) $q_j = p_i \cdot r_{ij}$, $q_k = p_i \cdot r_{ik}$ – calculation of partial transformants;

(iv) $N_j` = N_j + q_j$, $N_k` = N_k + q_k$ – finding the change of the CÆ-images ($CÆ_j$, $CÆ_k$) cardinals.

Note that the L-operator is a variant of the degenerate D-operator in which all r_{ij}, except one, are equal to zero.

F-operator (Fusion operator): (↑#, F, (n_i, ..., n_j), $r_{\cdot k}$) - a cardinal semantic operator of valency (W, V) = (w, 1), which assigns $r_{\cdot k}$ units of the transformant q_k to each *w*-tuple (n_i,... ,n_j) of CÆ-operands for the cardinal abstract entity $CÆ_k$.

A schematic representation of the F-operators $_2F$ and $_wF$ is shown in Fig.3.

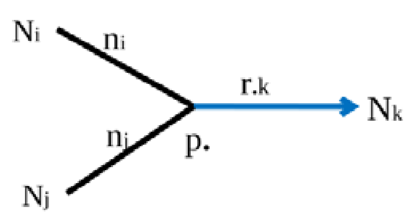


Figure 3.

When a $_2F$-operator acts on $CÆ_{i,j}$-operands, the following operations are performed:

(i) $p_i = \lfloor N_i/n_i \rfloor$, $p_j = \lfloor N_j/n_j \rfloor$ – calculation of partial carries;

(ii) $p = \min\{p_i, p_j\}$ – calculation of common carry;

(iii) $N_i` = N_i - p \cdot n_i$, $N_j` = N_j - p \cdot n_j$ – calculation of the remainders in $CÆ_i$, $CÆ_j$;

(iv) $q_k = p \cdot r_{\cdot k}$ – calculation of the transformant;

(v) $N_k` = N_k + q_k$ – finding the change of the $CÆ_k$-image cardinal.

Since the partial carries p_i, ..., p_j will be different, in the general case, the common carry must be determined from the condition of the existence of non-negative remainders in all CÆ-operands. This condition will be satisfied if we choose the minimal partial carry as the common carry *p*: $p_\cdot = \min \{p_i, ..., p_j\}$.

M-operator (Multi-operator): (↑#, M, (n_i,... ,n_j), ($r_{\cdot k}$,..., $r_{\cdot l}$)) - a cardinal semantic operator of valency (W, V) = (w, v), which assigns *v*-tuple conversion coefficients ($r_{\cdot k}$,..., $r_{\cdot l}$) of transformants to the *w*-tuple (n_i,... ,n_j) for CÆ-operands: $r_{\cdot k}$ units of k-transformant q_k for the cardinal abstract entity $CÆ_k$,..., and $r_{\cdot l}$ units of l-transformant q_l for the cardinal abstract entity $CÆ_l$.

A schematic representation of the M-operator ${}_2M_2$ is shown in Fig. 4.

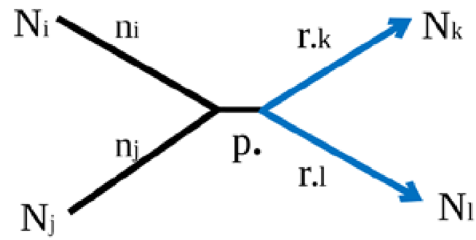


Figure 4.

When the ${}_2M_2$-operator acts on $CÆ_{i,j}$-operands, the following operations are performed:

(i) $p_i = \lfloor N_i/n_i \rfloor$, $p_j = \lfloor N_j/n_j \rfloor$ – calculation of partial carries;

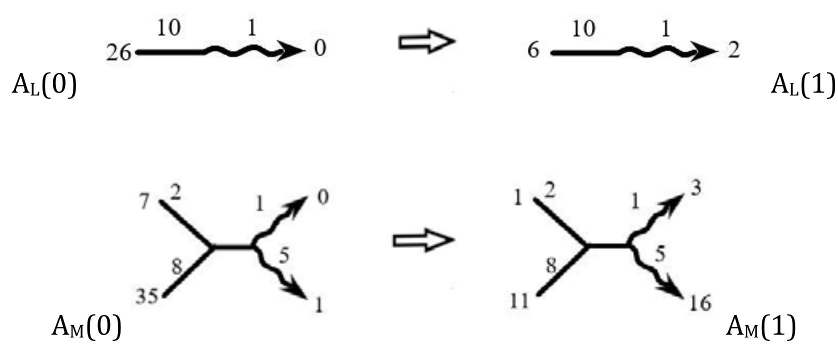
(ii) $p = \min\{p_i, p_j\}$ – calculation of common carry;

(iii) $N_i` = N_i - p \cdot n_i$, $N_j` = N_j - p \cdot n_j$ – calculation of the remainders in $CÆ_i$, $CÆ_j$;

(iv) $q_k = p \cdot r_{\cdot k}$, $q_l = p \cdot r_{\cdot l}$ – calculation of the partial transformants;

(v) $N_k` = N_k + q_k$, $N_l` = N_l + q_l$ – finding the changes of the $CÆ_{k,l}$-image cardinals.

Examples of cardinal semantic operators execution.



It is easy to see that any of the cardinal semantic operators considered above is a special case of the M-operator: $L \sim {}_1M_1$, $D_v \sim {}_1M_v$, ${}_wF \sim {}_wM_1$. However, for the construction of specific numeration systems and the analysis of cardinal semantic transforms in them, it is often more convenient to use such reduced forms of the M-operator.

These four cardinal semantic operators form the operator basis of any semantic numeration system.

NUMERATION SPACE. CARDINAL ABSTRACT OBJECT

To represent complex multistage semantic transformations, mono-operator transformations, as usual, are not enough. Let us introduce the concept of a *numeration space*, the elements of which are the numeration methods. By the *method of numeration* we mean a contextually conditioned method of transforming semantic units from a cardinal semantic multeity using cardinal semantic operators. Let us formalize the last statement with the concept of a *cardinal abstract object*.

Cardinal Abstract Object (CAO) is a collection of cardinal abstract entities connected in a certain topology by cardinal semantic operators. The signature of CAO_I:

$$\mathrm{Signt(CAO)} = (\mathbf{I}; \mathbf{CSM}; \mathbf{CSO}; \mathbf{STop}),$$

where **I** is the set of names denoting (naming) methods of numeration, **CSM** is the cardinal semantic multeity, **CSO** is the set of cardinal semantic operators, **STop** are the possible topologies of the semantic connectivity of cardinal abstract entities by cardinal semantic operators.

Numeration Space (NS) is an abstract space, the elements of which are cardinal abstract objects. A concrete CAO_I implements a specific method of numeration *I* in a numeration space.

The cardinal abstract entities included in the CAO that have an output, but do not have an input will be called *initial*, with input and output - *intermediate*, only with input - *final*, without input and output - *detached*. The main accepted assumption for the numeration methods (i.e. CAO) considered in this work is that any initial or intermediate cardinal abstract entity has a single output (associated with only one “perceiving” semantic operator) for an arbitrary (finite) number of inputs (transformants of other semantically consistent operators).

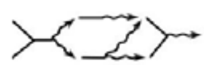
The concretizations of the CAO name, the composition of the cardinal semantic multeity, the type of operators and the topology of connectivity are determined by the *specification* of the cardinal abstract object.

Topology of Semantic Connectivity

The topology of semantic connectivity (STop) is determined by a given semantics of cardinal transformations and consists in connecting cardinal abstract entities from a cardinal semantic multeity by cardinal semantic operators of a given form.

The topology of semantic connectivity can be specified:

- descriptive/textually. For example, "serial connection of L-operators". Suitable for simple topology;
- in diagram form. For example:



- analytically: in the form of operator formulas of various types. For example, for the above graphical representation from left to right, top to bottom: $({}_2M_2|L, D_2|{}_2F)$;
- in tabular form.

The topology of semantic connectivity can be both "one operator"-type and "many operators"-type, both regular (for example, a 2-lattice, tree), and irregular, periodic or non-periodic, as well as cyclic. Linear topology can be specified recursively. For example, $(L)^m = L(L)^{m-1}$.

Thus, we can say that the positional numeration system in the traditional sense is a set of linearly connected cardinal abstract entities with bit semantics.

Cardinal Semantic Transformation. Multinumbers and Multicardinals

Let us agree to call a CSO "allowed" if the values of the cardinals of all its operands ensure the execution of the given operator.

Cardinal Semantic Transformation (CST) consists in executing, for a given CAO_I, all “allowed” cardinal semantic operators. A CST’ *step* will mean a single execution for a given CAO_I of all "allowed" cardinal semantic operators. The minimal sequence of CST’ steps, leading to steady values of all cardinals in CAO_I, will be called a *complete* CST, and the number of such steps will be called the *length* of CST.

A single implementation of the transformation will be called its *step*. Cardinal-semantic transformations can be both mono-operator and multi-operator. A mono-operator transformation is always one-step. Multi-operator transformations are usually multi-step.

The multiset of cardinals of all CÆs from CAO_I after an arbitrary step τ of a cardinal semantic transformations will be called the *multicardinal* of CAO_I of the step τ and denoted $\langle A_I(\tau)\rangle$ ($\langle A_I(\tau)\rangle = [N_i(\tau), N_j(\tau), \ldots, N_k(\tau)]$).

The multicardinal meaningfully characterizes only the “card-fullness” of CAO_I after each step of the cardinal semantic transformation, but in no way reflects the semantic aspect of the CST. The multicardinal of CAO_I before the first step of the transformation (CST) will be called the *initial* multicardinal $<A_I(0)>$, after a certain transformation step τ - the *intermediate* multicardinal $<A_I(\tau)>$, upon completion of transformations – the *final* multicardinal $<A_I(\omega)>$.

The holistic structural-cardinal representation of CAO_I after the τ-th step of the cardinal semantic transformation will be called the I-*multinumber* of the step τ (multinumber) and denoted by $A_I(\tau)$. By analogy with the multicardinal, before the first step of CST we will call a multinumber the *initial* multinumber $A_I(0)$, after a certain step (τ) of the transformation – the *intermediate* $A_I(\tau)$, upon completion of CST – the *final* multinumber $A_I(\omega)$.

We will assume that the multicardinal determines precisely the *meaning* of the CAO after the τ-th step of the cardinal semantic transformation, and the multinumber is its *sense*. Informally, a multinumber is a structured multicardinal, and a multicardinal is a de-structured multinumber.

Thus, a certain I-method of numeration (CAO_I) is a contextually determined complete cardinal semantic transformation of both multinumbers in the numeration space (NS) and the corresponding multicardinals in the cardinal semantic multeity (CSM).

Any number represented in one or another traditional positional numeration system is a multinumber, despite the absence of cardinal semantic operators in the numbers. This is due to both the linearity of the traditionally used CSOs, and the insignificant radix variability (negotiated separately), which allows you to write numbers into a string without distorting the meaning of the representation.

SEMANTIC NUMERATION SYSTEMS

Informally, the semantic numeration system (SNS) can be defined as a collection of homogeneous numeration methods.

By the *semantic numeration system* in the numeration space NS we mean its subspace SNS_φ with the given properties determined by the classification features. Here φ is the name-identifier of the semantic numeration system, due to a set of classification features.

We propose the following *classification of Semantic Numeration Systems*:

(1) by influence on the operand cardinal: transforming or preserving;

(2) by type of uncertainty: deterministic, stochastic and fuzzy;

(3) by kind of transformation: radix-multiplicity, radix-excess value; radix-excess fact;
arbitrary function;

(4) by kinds of number systems used in the numeration system: natural, integer or
rational numbers;

(5) by controllability: autonomous or controlled;

(6) by variability of the operator parameters (radices and conversion rates): homogeneous (the same for all operators) or heterogeneous (different for different operators);

(7) by topology of operators’ connectivity: linear (with not only L-operators), tree-like,
lattice, cyclic, amorphous or of a special form. Regular structures can be either isotropic or anisotropic, and the latter can be homogeneous or heterogeneous.

Thus, most of the generally accepted "numeration systems", for example, binary or decimal, will hardly need to be renamed. Within the framework of the above classification, they are particular (with n=2 and n=10, respectively) methods of numeration of the *transforming*, *deterministic*, *radix-multiplicity, natural, autonomous, linear*, *homogeneous* semantic numeration system.

FUZZY CARDINAL SEMANTIC TRANSFORMATIONS

In practice, one of the most crucial and common aspects of numerical data is its uncertainty. At present, mathematical theories that are considered to treat uncertainty are interval analysis [9], probability theory and the theory of fuzzy sets [10].

This state of affairs necessitates the development of tools for representing and processing uncertainty numerical information, in particular, in the field of semantic numeration systems. Since cardinal semantic transformations in semantic operators are basic for any SNS, we will focus on one-step transformations.

In principle, the following variances of the initial fuzziness are possible:

(i) "Fuzzy input, crisp operator". In this case, at least, one CÆ-operand has a fuzzy cardinal: card($CÆ_i$) = $\widetilde{N}_i$.

(ii) "Crisp input, fuzzy operator". All the input cardinals card(N_i), ..., card(N_j) are crisp. Some of the parameters (n, r) of CSO, or all of them, are fuzzy.

(iii) "Fuzzy input, fuzzy operator".

If a cardinal semantic operator is fuzzy, there are three possible ways of its performance: 1) when some radices are fuzzy with crisp conversion rates; 2) when radices are crisp with some fuzzy conversion rates; 3) when both the radices and the conversion rates are fuzzy.

Let $\tilde{A}$ be a continuous triangular fuzzy number and its membership function $\mu_{\tilde{A}}(x)$ is defined as [10]:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0, & x < \underline{a}, \\ (x - a)/(a - a), a & \le x \le a, \\ (\bar{a} - x)/(\bar{a} - a), a & < x \le \bar{a}, \\ 0, & \mathrm{x} > \bar{a}, \end{cases}$$

where x is an element of a support X, ($x \in$ X).

So each triangular fuzzy number $\tilde{A}$ may be defined as a triple $\tilde{A} = (a, a, \bar{a})$, where a is a lower bound, a is a mode, and $\bar{a}$ is an upper bound of the fuzzy number $\tilde{A}$.

All arithmetic operations on triangular fuzzy numbers are performed according to the rules of interval arithmetic [9, 10]. A feature of the cardinal semantic transformation in the radix-multiplicity kind of operators is the presence of the floor function $\lfloor * \rfloor$. So, we must introduce a rule for finding the floor function of both a fuzzy argument and a complex argument-operation on fuzzy numbers.

We assume that *the floor function of a triangular fuzzy number is a triangular fuzzy number, each component of which is the floor function of the corresponding component of the original number*.

Then,

$$\lfloor \tilde{A} \rfloor = \lfloor (\underline{a}, a, \bar{a}) \rfloor = (\lfloor \underline{a} \rfloor, \lfloor a \rfloor, \lfloor \bar{a} \rfloor), \forall \tilde{A} = (a, a, \bar{a}).$$

Respectively, for the floor function of the complex argument-operation on fuzzy numbers, for example, $\tilde{A} = (a, a, \bar{a}), \tilde{n} = (n, n, \bar{n})$ we can get the following expression:

$$\lfloor \tilde{A}/\tilde{n} \rfloor = \lfloor (\underline{a}, a, \bar{a}) / (n, n, \bar{n}) \rfloor = (\lfloor \underline{a}/\bar{n} \rfloor, \lfloor a/n \rfloor, \lfloor \bar{a}/\underline{n} \rfloor).$$

Similarly, expressions for more a complex argument-operation on fuzzy numbers can be obtained.

Thus, *fuzzy cardinal semantic transformation* is a cardinal semantic transformation performed according to the rules of fuzzy set arithmetic, fuzzy floor function and fuzzy common carry formation.

THE PRINCIPLE OF FUZZY COMMON CARRY FORMATION IN $_W F$ AND $_W M_V$ OPERATORS

Unlike the comparison of two fuzzy numbers and selection of a fuzzy number that is less than the other, in cardinal semantic operators $_W F$ and $_W M_V$ the fuzzy common carry has to be 'synthesized' based on a set of fuzzy partial carries $(\tilde{p}_i, \dots, \tilde{p}_j)$. Forming, but not choosing!

The concept of common carry calculation lies in its formation from the partial carries of each CÆ-operand of CSO, with guarantee that the cardinals of the remainders are zeroes or natural numbers in each CÆ of operands. In case that partial carries are represented as triangular fuzzy numbers, the following *"min-method"* of forming the fuzzy common carry is proposed.

The fuzzy common carry $\tilde{p}_. = (\underline{p}_., p_., \bar{p}_.)$ is formed from fuzzy partial carries $(\tilde{p}_i, \dots, \tilde{p}_j)$ so that its lower bound, mode and upper bound are defined as the minimum value of the corresponding bounds and the modes of the fuzzy partial carries:

$$\underline{p}_. = \min(\underline{p}_i, \dots, \underline{p}_j),$$
$$p_. = \min(p_i, \dots, p_j),$$
$$\bar{p}_. = \min(\bar{p}_i, \dots, \bar{p}_j).$$

Linear ordering of the components: $\underline{a} \leq a \leq \bar{a}$ for any $\tilde{A} = (\underline{a}, a, \overline{a})$ ensures the uniqueness of the solution.

An example of cardinal semantic transformation of fuzzy initial data by crisp $_3$F-operator is shown in fig. 5-7.

Let the initial cardinals be represented as triangular fuzzy numbers $\widetilde{N}_i = (9, 31, 45)$, $\widetilde{N}_j = (38, 73, 80)$, $\widetilde{N}_k = (29, 50, 101)$. The crisp $_3$F-operator has radices $n_i = 5$, $n_j = 10$, $n_k = 10$ and the conversion rate $r_{.l} = 2$ (fig.5).

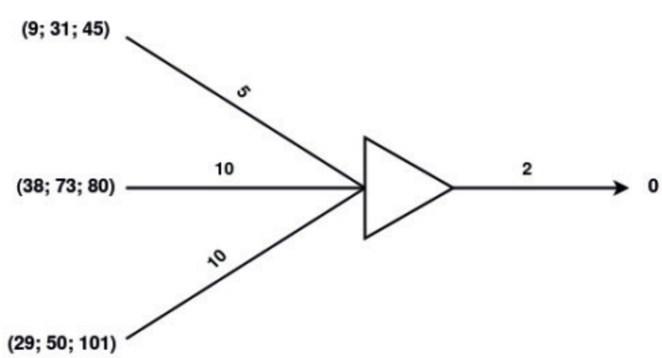


Figure 5.

Fuzzy partial carries obtained by the method $\tilde{p}_i = \lfloor \widetilde{N}_i / n_i \rfloor$, ..., and the fuzzy common carry $\tilde{p}_. = (1, 5, 8)$ formed by the "min-method" are shown in fig.6.

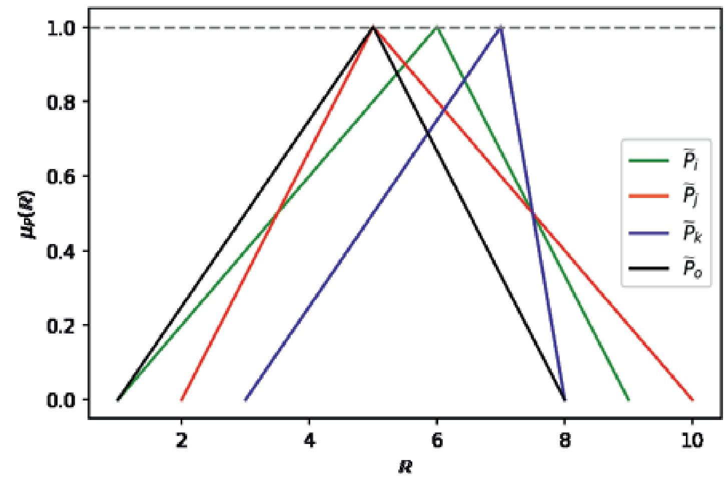


Figure 6.

The result of the fuzzy cardinal semantic transformation (fig.7) consists of CÆl-image fuzzy cardinal $\widetilde{N}_l = (2, 10, 16)$ and the corresponding fuzzy remainders.

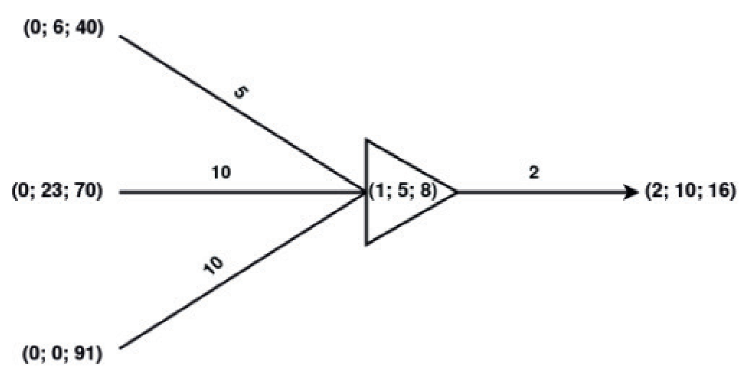


Figure 7.

CONCLUSION

The undoubted advantages of semantic numeration systems in comparison with existing positional ones are:

- a broader view on numeration systems as a way of representing not only numbers, but also sets, multisets, etc.;
- a variety of structures for representing numerical data within the framework of a single numeration method from positional to position-structured;
- an emphasis on the difference between meaning and sense in numeration systems. Sense as the entanglement of meanings and structure.

The extension of SNS to fuzzy semantic numeration systems makes it possible to take into account in the representation of numerical data such types of uncertainty as interval and fuzzy. This makes fuzzy semantic numeration systems indispensable in applied artificial intelligence systems.

It seems important to research the possibility of creating semantic numeration systems with a probabilistic type of uncertainty. For applications, it is necessary to explore a number of "mix directions" such as combination of different types of uncertainty in one CSO (CAO), for example, fuzzy and probabilistic, as well as the possibility of combining discrete and continuous fuzzy numbers in one cardinal semantic operator.

References

[1] Fraenkel A. (1985). Systems of Numeration. *Amer. Math. Monthly*, V.92, 2, 105-114.

[2] Rigo M. (ed.) (2014). *Formal Languages, Automata and Numeration Systems 2. Applications to Recognizability and Decidability*. Wiley.

[3] Dimitrov V., Jullien G., Muscedere R. (2012). *Multiply-Base Number System: Theory and Application*. CRC Press.

[4] Chunikhin A. (2012). *Introduction to Multidimensional Number Systems. Theoretical Foundations and Applications*, LAP LAMBERT Academic Publishing, (in Russian).

[5] Chunikhin A. (2018). Multidimensional Numbers and Semantic Numeration Systems: Theoretical Foundation and Application. *JITA – Journal of Information Technology and Applications*, PanEuropien University APEIRON, Banja Luka Republika Srpska, Bosna i Hercegovina, JITA 8, 2:49-53; doi: 10.7251/JIT1802049C.

[6] Chunikhin A. (2021). Fundamentals of semantic numeration systems. Can the context be calculated? *arXiv:2102.09949v2*. doi:https://doi.org/10.48550/arXiv.2102.09949.

[7] Chunikhin A. (2022). Semantic numeration systems as information tools. *Proceedings*, 81, 23. doi:https://doi.org/10.3390/proceedings2022081023.

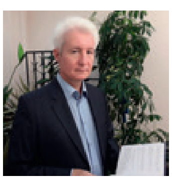
[8] Burgin M., Dodig-Crnkovic G. (2020). Prolegomena to an Operator Theory of Computation. *Informations*, 11, 349; doi: 10.3390/info11070349.

[9] Moor R.E. (1979). *Methods and Application of Interval Analysis*. SIAM, Philadelphia. doi:https://doi.org/10.1137/1.9781611970906.

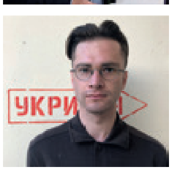
[10] Piegat A. (2001). *Fuzzy Modeling and Control*. Physica Heidelberg. doi:https://doi.org/10.1007/978-3-7908-1824-6.

Received: March 2, 2022
Accepted: May 19, 2022

About the authors



Alexander Ju. Chunikhin, Candidate in Engineering, received the diploma of engineer in radioelectronics from Higher School of Military Aviation Engineering (Kiev, URSS) in 1983. He received the PhD (Candidate in Engineering) degree from Higher School of Military Aviation Engineering (Kiev, Ukraine) in 1991. Currently he works as a Senior Researcher of Palladin Institute of Biochemistry (The National Academy of Sciences of Ukraine). His research interests include complex systems, Petri nets, number systems and semantic numeration systems. He published more than 100 scientific papers, two monographs.



Vadym Zhytniuk is a master's student at Kyiv National University. He is currently enrolled in an internship at the Bogomoletz Institute of Physiology in the Department of Molecular Biophysics and occupies a position of a lab assistant at the Palladin Institute of Biochemistry in the Department of Muscle Biochemistry. His interests include mathematical aspects of complex biological systems research.

For citation

Alexander Yu. Chunikhin, Vadym Zhytniuk, Semantic Numeration Systems as Information Tools for Fuzzy Data Processing, *JITA – Journal of Information Technology and Applications, Banja Luka*, Pan-Europien University APEIRON, Banja Luka, Republika Srpska, Bosna i Hercegovina, JITA 12(2022) 1:14-25, (UDC: 004.6.056.55:519.852), (DOI: 10.7251/JIT2201014C), Volume 12, Number 1, Banja Luka, June (1-64), ISSN 2232-9625 (print), ISSN 2233-0194 (online), UDC 004

PREPRNA: AN INTEGRATIVE TOOL FOR ILLUMINA RNASEQ DATA FILTERING

Dragana Dudić[1], Bojana Banović Đeri[2], Željko Stanković[3], Zoran Ž. Avramović[3]

[1]*Faculty of Informatics and Computer Science, University Union Nikola Tesla, Belgrade, Serbia,*
ddudic@unionnikolatesla.edu.rs

[2]*Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Serbia,*
bojanabanovic@imgge.bg.ac.rs

[3]*Pan-European University APEIRON, Banja Luka, Bosnia and Herzegovina,*
{zeljko.z.stankovic, zoran.z.avramovic}@apeiron-edu.eu

Contribution to the State of the Art

https://doi.org/10.7251/JIT2201026D UDC: 547.963.3:577.113

Abstract: The vast amount of currently available transcriptome sequences is comprised of Illumina RNAseq data. Usually, publicly available datasets are provided as raw data and preparing them for the downstream NGS analysis is the first step required. Such preprocessing step, besides the evaluation of the quality of the raw data, includes data filtering, in order to provide high quality results of the downstream analysis. Existing tools for NGS data filtering are either too general or incomplete for the Illumina RNAseq filtering task, which is why a new tool for this endeavor was needed. We present prepRNA, a novel tool intended for Illumina RNAseq data filtering, which was designed as a comprehensive and user-friendly wrapper tool with possibility of further upgrading with a quality control option.

Keywords: RNAseq, data filtering, data preprocessing, NGS data, Illumina.

INTRODUCTION

In all eukaryotes genome content is transcribed from DNA molecule to RNA molecule by complex biological process. Entirety of RNA molecules, which may arise in the transcription process, make a complete transcriptome of that organism. Total eukaryotic transcriptome is comprised of two different classes of RNA molecules: 1) coding RNAs - messenger RNAs (mRNAs) and 2) non-coding RNAs – ribosomal RNAs (rRNAs), transport RNAs (tRNAs), short RNAs (miRNAs, piRNAs, siRNAs, snRNAs, snoRNAs) and long non-coding RNAs (pRNAs, eRNAs, gsRNAs, lincRNA, NAT), with different functions. Moreover, total transcriptome has a variable content and structure, and its composition directly depends on internal and external conditions, enabling real-time adaptation of an organism to developmental stage and surrounding environment. The main causes of variability in transcriptome content in eukaryotes include alternative splicing, inclusion, exclusion and up- or down-regulation of expression of individual genes via their regulation pathways.

RNA sequencing (RNAseq) represents determination of the sequence of nucleotides in total transcriptome of one organism. Since tRNAs and rRNAs are considerably more abundant than the rest of RNAs, they are usually physically removed during NGS library preparation in order not to mask the rest of the transcriptome. Further, total transcriptome contains numerous different mRNA transcripts at different levels of intron processing. Determining complete RNA content and presence of individual transcripts in RNA, their number and possible nucleotide changes in different organisms, different tissues of one organism, different stages of development and different environmental conditions are important for gene expression determination, leading to insights in potential genes' roles and pathways of regulation as well as in phylogenetic analysis and evolution of certain molecular pathways, gene families, etc.

For the last 40 years, determining total RNA content and content of individual RNA classes and transcripts has been an active issue, during which several sequencing technologies were developed. These new sequencing technologies included: Roche 454, SOLiD, Ion Torrent, Illumina, PacBio and Oxford Nanopore platforms, based on different approaches to determine precise sequence with deep coverage. Increased availability of such Next Generation Sequencing (NGS) technology provided the vast amount of RNA-seq data, since it became favored for transcriptome analysis due to higher coverage and resolution when compared to Sanger and microarray sequencing [1].

Overall insight in publicly available NCBI GEO datasets revealed that among 161,145 published RNAseq datasets (date of search March 17th 2022), 255 originated from ION Torrent technology, while 160,890 originated from Illumina, implying that Illumina platform is the most popular choice for RNA-seq analysis among researchers.

Filtering sequenced data

First step in any NGS analysis is to preprocess the NGS data: to check the quality of the raw data and, according to results of the quality control, to remove the adapters' sequences, contaminating sequences and to omit low quality portions of the NGS dataset (NGS data filtering). Thus, in the Illumina RNAseq domain, necessary data filtering tasks include [2]: contaminant removal in a narrow sense, adapter removal and trimming of low-quality nucleotides.

One of the steps in the process of preparing a sample for NGS sequencing includes adding short sequences of known and constant nucleotide content, called adapters, at the end of every fragment. Although adapters should not be present in the resulting raw set of NGS reads, sometimes, when the length of sequenced molecules vary, as commonly happens when working with total transcriptome (due to the nature of transcriptome composition), adapters are retained at 3' end of a sequence. Their presence can be noticed during the NGS quality control, where several checkpoints indicate abundant presence of the adapters. Even if the quality control checkpoints do not show the presence of the adapters, it still does not exclude their presence, because the quality control tool may not report adapters if they are not abundant enough to reach a sufficient extent in a randomized sample. Nevertheless, adapters should be removed from the NGS dataset, because their presence degrades the quality of the obtained results. Usually, general tools for trimming/preprocessing of NGS data, like Trimmomatic [7], Cutadapt [8] and AfterQC [9], are used for the adapter removal task, but there are also other, more specialized tools, like Scythe.

Contaminants in a narrow sense include all sequences originating from different organisms than the one that was sequenced and as such should be removed from the sequenced data. The first insight in the level of contamination is given by the quality control checkpoints, but in order to get all information further exploration is needed. Contaminant organisms can be detected by sampling 500-1,000 sequences from the RNAseq dataset and searching against the BLAST database. If contaminant organisms are detected, level of contamination can be determined by specific tools for this task, like VecScreen [3] and FastQScreen [4] or with broader purpose tools that have this functionality, like DeconSeq [5]. Contamination removal may be done by using specialized tools like BioBloom or by tools with broader scope like FastqPuri [10] and DeconSeq or even by mapping the NGS data to the reference genomes of the sources of the contamination.

All currently available NGS platforms have certain technological limitations and the presence of some low-quality nucleotide sequences is expected in the sequence dataset. Illumina NGS sequencing uses approach called sequencing-by-synthesis, which is error-prone in its nature and may be the reason of occasional low-quality nucleotides to occur. Usually, in order to deal with these parts of the NGS sequence, the low-quality base trimming is used. Low quality base trimming is a necessary step in a filtering process and it is included in all general tools for filtering/preprocessing of NGS reads. Initial checkpoints of NGS quality control indicate position and level of low-quality bases, which should be used as guidelines for the lower quality boundary in the NGS reads and decision on their removal.

Related work

There are many available tools for filtering NGS data, but among them there are very few designed for filtering RNAseq data, and none of them is specialized for Illumina data. That is why, until now, the only option was to use more general tools for RNAseq data filtering tasks, like Trimmomatic, Cutadapt and AfterQC. Trimmomatic is a tool for filtering Illumina NGS data, Cutadapt is a fully general NGS filtering data tool, while AfterQC is a tool intended for NGS data preprocessing with partial quality control support.

From the perspective of Illumina RNAseq filtering approach, all three tools have similar functionalities: they all have features for quality trimming, adapter removal and removing of the reads below specified length. While AfterQC tool removes adapters automatically, Trimmomatic and Cutadapt, as general tools, have to be provided with the adapters' sequences specific for the used platform. Among two later general tools the difference is in a way of forwarding of adapters' sequences: Trimmomatic expects an input file with adapters' sequences and Cutadapt accepts adapters' sequences as command line argument. Users may have troubles with providing adapter sequences if they are not familiar with sequencing technology and how to find the list of used adapters. Trimmomatic, as a tool designed for Illumina NGS data, offers several files with usual Illumina adapters, but still, the user may be confused which one to choose if they do not have information about the sequencing process. Also, all three tools have a similar disadvantage by missing the feature for contaminant removal in a narrow sense, because this feature is needed to facilitate the further downstream analysis.

Besides general NGS tools, there are also tools specialized for RNAseq data, like FastqPuri and RNA-QC-Chain [11]. RNA-QC-Chain is a tool for RNAseq data preprocessing with partial quality control support and no contaminant removal in a narrow sense. According to [10], FastqPuri is a tool for Illumina RNAseq data preprocessing with partial quality control support, but it lacks user instructions, documentation and support. Also, tool installation failed on several Unix machines because of unsupported R packages.

Due to all above mentioned issues, we created prepRNA, as a user-friendly tool meant for Illumina RNAseq complete filtering, which promotes reusing and repurposing instead of reinventing, through combining the best solutions from the existing sources.

Implementation

prepRNA is a command-line Unix wrapper tool built around Trimmomatic and Bowtie tools, which enables comprehensive and user-friendly complete filtering of Illumina RNAseq data. It was intended to be used after finished quality control of RNAseq data and relies on the information on determined level of contamination. The general architecture of a prepRNA tool is presented on Figure 1.

prepRNA consists of three modules: contaminant-index-download module, contaminant-indexing module and filtering module. All modules are easy-to-use with a clear and concise user manual given under --help option. Contaminant-index-download module represents an optional module that enables downloading reference genomes for commonly rep-

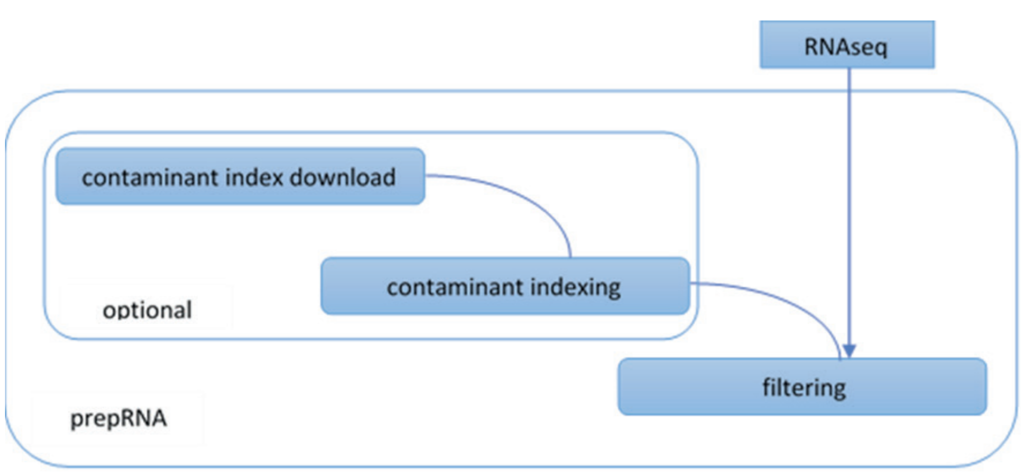


Figure 1. prepRNA tool architecture

resented Illumina RNAseq contaminants in a narrow sense, which include: PhiX bacteriophage as control viral DNA, vectors like plasmid, phage, cosmid, BAC, PAC, YAC and transposable elements from the cloning host, which is usually *Escherichia coli* or yeast and human as the NGS library and NGS platform handler. Options offered in this module are straight-forward and every listed contaminant has a separate option which invokes its download. Tool creates a directory inside the current directory where downloaded DNA sequences are stored. After downloading required contaminant reference genomes, they are automatically indexed in order to enable decontamination of Illumina RNAseq data. This is performed within the contaminant-indexing module, which was built as the wrapper around bowtie-build functionality for indexing sequencing data of the Bowtie tool. Bowtie tool is a well-known and broadly used tool for mapping of DNAseq data, but in this case, it was used for indexing and mapping of RNAseq data of selected contaminant organisms, because no splice-awareness is needed for decontamination of RNAseq data. Once created, indexes are stored together with contaminant sequences and could easily be reused in other sequencing projects. Since contaminant-indexing module depends on the contaminant-index-download module, they share the same options.

The filtering module was built as a wrapper around Trimmomatic and Bowtie tools, offering necessary and sufficient options for preprocessing of Illumina RNAseq data: contaminant removal in narrow sense, adapter removal and quality trimming. For all of the aforementioned tasks, a separate option is offered. Contaminant removal option expects value as a string that contains first letters of contaminants in any order. Quality trimming option expects integer value between 1 and 40, covering the full range of quality values available in Illumina RNAseq data. Adapter removal is fully automated step and it has no value, because the domain of interest is specialized to Illumina RNAseq. Additionally, filtering of extremely short nucleotide sequences was set to be done automatically, while duplicate removal and removal of the content bias at the 5' end were omitted, as not beneficial for the further analysis of the Illumina RNAseq data [2].

Regarding the steps of adapter removal, quality trimming or adapter removal together with quality trimming, prepRNA was created as a wrapper tool around Trimmomatic. Trimmomatic is a widely used tool for filtering Illumina sequencing data characterized by good performance, but not as user-friendly for people not having good computer skills. Moreover, Trimmomatic lacks some important functionalities necessary for preprocessing of RNAseq data, like removal of contaminants in a narrow sense, which is why we added this functionality by using the mapping method with the Bowtie tool, through which RNAseq data are mapped to the selected downloaded and indexed contaminant reference genome. Only unmapped data are then processed and forwarded for further RNAseq filtering or offered for output in adequate format, depending on the task. Also, in prepRNA tool the parameters used in adapter removal, quality trimming or adapter removal together with quality trimming are automatically set to the values which showed the best performance for RNAseq data globally. In that way, prepRNA is easy to use by non-computer specialized users, like biologists, physicists and chemists.

Evaluation

There is a lack of tools for raw Illumina RNAseq data filtering, which is why more general filtering tools are being used for this task, leading to different difficulties for non-computer specialized users. Although most of general tools give good results for selected tasks, none of them offer a complete list of filtering tasks needed for the Illumina RNAseq data. Unlike similar existing tools, prepRNA provides complete and comprehensive filtering of Illumina RNAseq data. Comparison of five existing filtering tools with prepRNA tool is given in the Table 1.

In the matter of domain, among six compared tools, three tools were designed for RNAseq data and two tools were adapted for Illumina data, but only prepRNA was designed for both Illumina and RNAseq data. Regarding user friendliness and easiness for installation three out of six tools met both criteria: Cutadapt, AfterQC and prepRNA, but among them only prepRNA was designed for RNAseq and Illumina data. Features comparison showed that among six tools only two, prepRNA and FastqPuri, provided all the features needed for preprocessing of Illumina RNAseq data, but FastqPuri did not meet two other important criteria of being easy to

Table 1. *Comparison of tools for filtering RNAseq data. QC – quality control, CR -contaminant removal, AR- adapter removal, QF – quality filtering*

Tool	prepRNA	Trimmomatic	Cutadapt	FastqPuri	AfterQC	RNA-QC-Chain
Language	C, java	java	C,Python	C, R	C,Python	C++
Mode	SE, PE	SE, PE	SE, PE	SE, PE	PE	SE
Format	fq.gz	fq*	fq, fa, gz	fq*	fq	fq,fa
RNAseq	+	-	-	+	-	+
Illumina	+	+	-	-	-	-
easy to install	+	+	+	-	+	+
easy to use	+	-	+	-	+	-
QC	-	-	-	+/-	+/-	+/-
CR	+	-	-	+	-	-
AR	+	+	+	+	+	+
QF	+	+	+	+	+	+

install and easy to use, while prepRNA did. Regarding this last consideration one should have in mind that aforementioned tools are designed for Unix-like systems and mostly used on server machines, where regular users do not have permissions for installing software anywhere but the home directory. This makes installing dependencies difficult, which may disable functionalities of a tool. That is why it is more convenient to use tools developed on programming languages that are built in the system, because compatibility and functionality are guaranteed in that case.

Considering tools' availability to be used with two reading modes, single-end (SE) and paired-end (PE), most of aforementioned tools, including prepRNA, supported both ways. This feature is important because Illumina PE sequencing is widely used RNA sequencing method and supporting PE reading in RNAseq filtering is necessary. In PE reading, obtained sequences are more reliable because sequencer reads every fragment twice, rather than in SE reading, where fragments are read once. In respect of tools support of sequencing formats, it was essential to support the widely used compressed format for Illumina RNAseq data, fastq.gz (fq.gz) format, which stand for all aforementioned tools, including prepRNA. Finally, RNAseq data filtering resides on the results of quality control, which makes convenient for RNAseq quality control and RNAseq data filtering to be combined in one tool. Even though three out of six tested tools - FastqPuri, AfterQC and RNA-QC-Chain tried to implement this task, none of them provided comprehensive quality control needed for RNAseq data. Thus, prepRNA did not tried to implement it, but rather to lean on output of existing comprehensive quality control tools, like FastQC.

Conclusion

Existing tools that can be used for Illumina RNAseq filtering tasks lack some functionalities necessary for the completeness of the filtering process. Thus, we developed prepRNA, a user-friendly and comprehensive tool for NGS data filtering designed especially for Illumina RNAseq data. prepRNA is maintained, easy-to-install and easy-to-use tool with clear user manual and installation instructions. This tool was designed as a wrapper tool, combining two broadly used tools in the bioinformatics community, with an aim to fulfill all the requirements of the Illumina RNAseq data filtering process. It provides contaminant removal, adapter removal and quality trimming steps, for which it uses parameters, determined to be the best from global RNAseq practice, as default ones. Moreover, prepRNA can be extended with quality control functionality in order to address the whole process of Illumina RNAseq data preprocessing. Finally, prepRNA takes user's input in short and unambiguous form expecting only RNA sequences files and desired filtering options provided as an input, making it highly desirable among non-computer specialized users.

References

[1] Kukurba, K. R., & Montgomery, S. B. (2015). RNA Sequencing and Analysis. *Cold Spring Harbor protocols*, *2015*(11), 951–969. https://doi.org/10.1101/pdb.top084970

[2] Dudić, D., Đeri, B. B., Pajić, V., & Pavlović-Lažetić, G. (2021). Demystification of RNAseq Quality Control. JITA, 22(2), 73-86.

[3] Schäffer, A.A., Nawrocki, E.P., Choi, Y., Kitts, P.A., Karsch-Mizrachi, I., McVeigh, R. (2018) Vec-Screen plus taxonomy: imposing a tax(onomy) increase on vector contamination screening. Bioinformatics 34(5), 755–759

[4] Wingett, S. W., & Andrews, S. (2018). FastQ Screen: A tool for multi-genome mapping and quality control. F1000Research, 7, 1338. https://doi.org/10.12688/f1000research.15931.2

[5] R. Schmieder, R. Edwards (2011). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. PLoS One, 6: e17288, doi: https://doi.org/10.1371/journal.pone.0017288.

[6] Pérez-Rubio, P., Lottaz, C. & Engelmann, J.C. (2019) FastqPuri: high-performance preprocessing of RNA-seq data. BMC Bioinformatics 20, 226 . https://doi.org/10.1186/s12859-019-2799-0

[7] Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics, btu170.

[8] Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal, 17(1), pp. 10-12. doi:https://doi.org/10.14806/ej.17.1.200

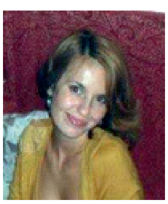
[9] Chen, S., Huang, T., Zhou, Y. et al. (2017) AfterQC: automatic filtering, trimming, error removing and quality control for fastq data. BMC Bioinformatics 18, 80 . https://doi.org/10.1186/s12859-017-1469-3

[10] Pérez-Rubio, P., Lottaz, C. & Engelmann, J.C. (2019) FastqPuri: high-performance preprocessing of RNA-seq data. BMC Bioinformatics 20, 226 . https://doi.org/10.1186/s12859-019-2799-0

[11] Zhou, Q., Su, X., Jing, G. et al. (2018) RNA-QC-chain: comprehensive and fast quality control for RNA-Seq data. BMC Genomics 19, 144 . https://doi.org/10.1186/s12864-018-4503-6

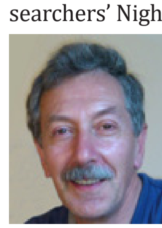
Received: December 12, 2021
Accepted: April 1, 2022

About the authors



Dragana Dudić is an Assistant at the Faculty of Informatics and Computer Science at the University Union Nikola Tesla for several Computer Science subjects. She obtained BSc and MSc in Computer Science at the Faculty of Mathematics, University of Belgrade where she is finishing her PhD in the field of Bioinformatics. Dragana is interested in research of repetitive sequences, especially human and plant transcriptomics.

Bojana Banović Đeri is a Associate Research Professor in the Laboratory for Plant Molecular Biology in the Institute of Molecular Genetics and Genetic Engineering, University of Belgrade. In 2013. obtained PhD in Molecular biology at the Faculty of Biology, University of Belgrade in the field of plants' molecular biology, genetics and proteomics (characterization of buckwheat heteromorphic self-incompatibility), with later specialization in bioinformatics (until 2021: nine international trainings and programme "Bioinformatics for biologists" at the School of Computing, Union University, Belgrade). In 2017 participated in the US expert exchange programme „Open World" for food safety and security. Her scientific interests primarily includes plants' response to stress, single stranded RNA viruses and food safety. Author of 19 research papers (https://orcid.org/0000-0001-7663-1878) cited more than 135 times (Google Scholar). Coordinator of one international and one national project and participant in 14 projects (national and international). Lecturer at Molecular Biology module for PhD students, Faculty of Biology, University of Belgrade. Mentor of one PhD thesis and three master thesis. Active as a member of several scientific societies, reviewer of international scientific papers and projects, organizer of several conferences and workshops. Engaged in science promotion (Researchers' Night, Plant Fascination Day, etc.).



Željko Stanković was born in Belgrade. He finished primary and secondary school in his hometown. He received his higher education in Cleveland, Ohio, USA, where he graduated in 1981. In 2006, he defended his master's thesis at the University of Novi Sad. He defended his doctoral dissertation at Singidunum University in 2010. He has been programming since 1984. making programs for his first computer Commodore 64. Robotics and bioengineering are a long-standing field of work and interest.

Zoran Ž. Avramović was born in Serbia, on September 10, 1953. He finished elementary school and high school in Loznica with great success. He was awarded several diplomas by Nikola Tesla and Mihailo Petrović Alas. He graduated on time at the University of Belgrade - Faculty of Electrical Engineering, with an average grade of 9.72 in five-year studies. He received his master's degree at that faculty (all excellent grades, exams and master's degrees), and then obtained a doctorate in technical sciences (in 1988). As an excellent student of the University, he had the right and at the same time studied mathematics at the Faculty of Mathematics in Belgrade. He was the champion of Serbia in mathematics ("first prize") and Yugoslavia in electrical engineering ("gold medal")

For citation

Dragana Dudić, Bojana Banović Đeri, Željko Stanković, Zoran Ž. Avramović, prepRNA: an integrative tool for Illumina RNAseq data filtering, *JITA – Journal of Information Technology and Applications, Banja Luka*, Pan-Europien University APEIRON, Banja Luka, Republika Srpska, Bosna i Hercegovina, JITA 12(2022) 1:26-31, (UDC: 547.963.3:577.113), (DOI: 10.7251/JIT2201026D), Volume 12, Number 1, Banja Luka, June (1-64), ISSN 2232-9625 (print), ISSN 2233-0194 (online), UDC 004

Evaulation of Homomorphic Encryption Implementation on IoT Device

Goran Đorđević[1], Milan Marković[2], Pavle Vuletić[3]

[1]*AET Europe, Arnhem, The Netherlands, School of Electrical Engineering, Belgrade, Serbia, goran.djordjevic@aeteurope.com*

[2]Pan-European University Apeiron, Banja Luka, Serbian Republic, BiH, *milan.z.markovic@apeiron-edu.eu*

[3]School of Electrical Engineering, Belgrade, Serbia, *pavle.vuletic@etf.bg.ac.rs*

Contribution to the State of the Art

https://doi.org/10.7251/JIT2201032DJ UDC: 004.738.5:621.391

Abstract: An encryption scheme is homomorphic if it supports operations on encrypted data. Homomorphic encryption allows a device to perform arbitrary computations on encrypted data without user secret key. Recently it is introduced new homomorphic encryption schemes with improved performance that can be implemented in IoT device in production environments. The IoT concept encompasses devices, sensors, and services existing within an interconnected infrastructure with an efficient access to sample computational facilities. In this paper we evaluated features of exact arithmetic homomorphic encryption mechanisms: BFV and BGV and approximate homomorphic encryption scheme: CKKS. In the paper we measured performances of operations of homomorphic encryption schemes: BGV, BFV and CKKS that are implemented in Raspberry Pi 4 IoT device.

Keywords: Homomorphic encryption, Raspberry Pi IoT device, HE schemes performance.

Introduction

Traditional encryption schemes, both symmetric and asymmetric, were not designed to perform computations on the ciphertext in a way that would pass through the encryption to the underlying plaintext without using the secret key. The property would in many contexts be considered a vulnerability.

Homomorphic encryption differs from basic encryption methods in that it allows computation to be performed directly on encrypted data without requiring access to a secret key [1]. Homomorphic encryption schemes can be taken as a generalization of public key encryption mechanisms.

For long time the main issue of homomorphic encryption implementation was its low performance. Due to the issue the homomorphic encryption could not applied in production environments. However, in last few years it is introduced new homomorphic schemes with improved performance that make them suitable appliance in production systems. Additionally, it is proposed new optimization mechanisms of existing homomorphic operations that improved them performance and they can be used in production systems.

Fully Homomorphic Encryption (FHE) scheme is cryptographic mechanism that support arbitrary level of computations on ciphertext (addition, rotation, multiplication) not knowing the value of its secret key. Modern fully homomorphic encryption algorithms use complex mechanims on lattice structures and Ring-LWE (Ring Learning With Errors, RLWE) scheme [2].

According to latest researchs modern homomorphic encryption schemes are resistant to quantum computer attacks. There are no known mechanisms that would use the features of quantum computers to break homomorphic algorithms in polynomial time.

An example of cloud-based scenario with deployed homomorphic encryption operations is shown in Figure 1.

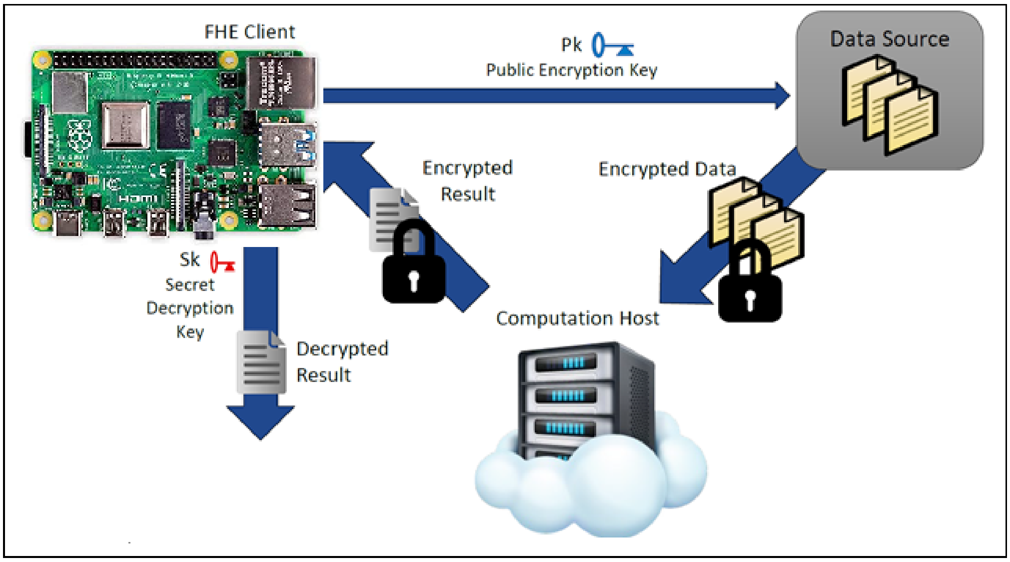


Figure 1: *Client-server homomorphic encryption scenario*

At server side data is processed on the server in the encrypted form. The results are remained in the encrypted form and they are sent back to the device who can decrypt the data and use the result.

In the work homomorphic encryption mechanisms are implemented in Raspberry Pi 4 Model B. The first generation of Raspberry (Pi 1) was released in the year 2012, that has two types of models namely model A and model B.

In this work we are estimated the features and performance of exact (BGV and BFV) and approximate (CKKS) homomorphic encryption schemes are discussing the conveniences and constraints of these schemes for use in the cloud-based systems. Performance assessment of the FHE schemes implemented in IoT device has not been much explored in the literature. The work will provide a better insight into the current state of the work on homomorphic encryption and its suitability for deployments in real IoT based systems.

The CKKS homomorphic scheme enables approximate homomorphic encryption for efficient processing of floating-point data. Using CKKS homomorphic scheme can be considered as efficient choice for homomorphic logistic regression and other machine learning tasks.

The paper is organized as follows. Section 'Related work' gives an overview of the related work in the field of the performance evaluation of the FHE schemes. The most important properties of homomorphic encryption and the classification of the homomorphic encryption schemes are presented in Section 'Properties of homomorphic encryption'. An application of Homomorphic Encryption: IoT (Internet of Things) case is shown in Section 'An example of appliance of homomorphic encryption'. The main features and description of modern HE schemes: BGV [3], BVF [4] and CKKS [5] are elaborated in Section 'Homomorphic schemes'. In Section ‚Experimental analysis' is shown results of experimental analysis. Conclusions are given in Section 'Conclusions'.

Related work

In article [6] is described only BGV Encryption Scheme for IoT Systems. In the work is not shown homomorphic encryption operations performance for other relevant modern schemes like CKKS and BFV.

In article [7] is analyzed and described test results only for CKKS approximate homomorphic encryption scheme. In the work is not described neither shown homomorphic encryption operations performance for other schemes like BGV and BFV.

In article [8] is described Fast Number Theoretic Transform operations for Ring-LWE on 8-bit AVR Embedded Processor but in the article is not shown performance of homomorphic encryptions schemes like CKKS, BFV and BGV.

Unlike the previous work, in this work we give experimental results for exact (BGV, BFV) and approximate (CKKS) homomorphic encryption schemes that are implemented in IoT device.

PROPERTIES OF HOMOMORPHIC ENCRYPTION

There are four main types of homomorphic schemes [9]:

- Partially Homomorphic Encryption (PHE). The partially homomorhpic encryption supports any number of operations (multiplication, addition), but it is limited to just one type of operation.
- Somewhat Homomorphic Encryption (SHE) allows both addition and multiplication, but it can perform a limited number of operations.
- Levelled Homomorphic Encryptions (LHE). This scheme can evaluate only functions with certain complexity (depth).
- Fully Homomorphic Encryption. The scheme allows any number of addition or multiplication operations. An Fully Homomorphic Encryption scheme can evaluate unbounded complexity (depth).

Homomorphic Encryption should support two main homomorphic operations [10]:

- Additive Homomorphic Encryption (Figure 2).
- Multiplicative Homomorphic Encryption.

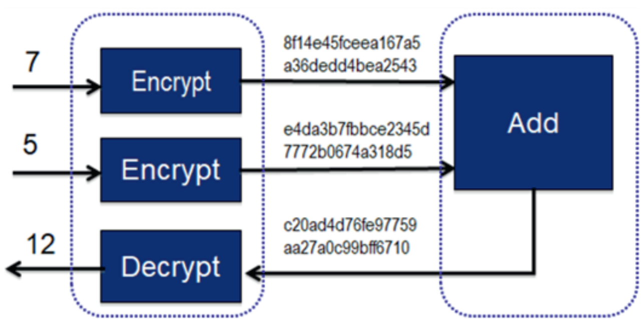


***Figure 2:** An example of Additive Homomorphic Encryption*

Homomorphic encryption is additive, if [11]: Enc $(m_1 + m_2)$ = Enc (m_1) + Enc (m_2); $\forall m_1, m_2 \in M$.

Homomorphic encryption is multiplicative, if [11]: Enc $(m_1 * m_2)$ = Enc (m_1) * Enc (m_2); $\forall m_1, m_2 \in M$.

The most popular classes of homomorphic schemes, given with their main properties (Figure 3), are:

- Boolean circuit (Fastest Homomorphic Encryption in the West (FHEW) [12] and Fast Fully Homomorphic Encryption over the Torus (TFHE) [13]):
  - Plaintext data are coded as bits;
  - Computations are performed by using Boolean circuits.
- Modular integer arithmetic (BGV, BFV):
  - Plaintext data are coded as integer modulo a plaintext;
  - Computations are expressed as integer modulo arithmetic.
- Approximate number arithmetic (CKKS):
  - Plaintext data are coded as real (or complex) numbers;
  - Computations are performed in a way similar to floating-point arithmetic but dealing with fixed-point numbers.

Modern homomorphic encryption schemes are based on usage of lattice cryptography with errors LWE [14]. Lattices have an important role in modern cryptography, especially in the context of the research on post-quantum cryptography. So far it is not reported in the literature fact that is claimed

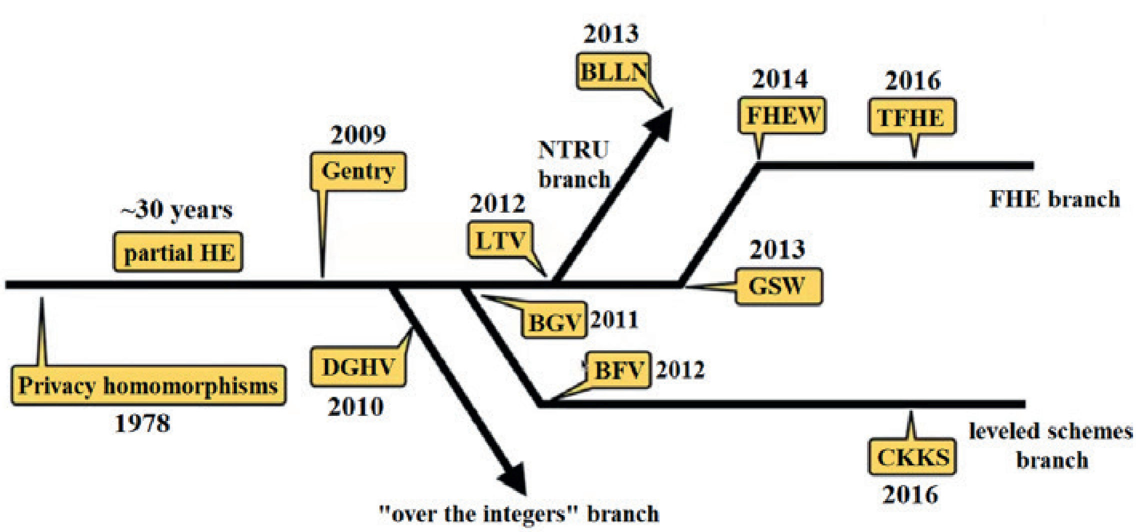


***Figure 3:** History of Homomorphic scheme*

that it can break lattice-based cryptographic algorithms using quantum computer algorithms.

The modern homomorphic encryption schemes apply structured lattices i.e. they use Ring-LWE mechanism [2]. The Ring-LWE reduces both important factors: computation time and key length.

The Ring-LWE implementation is based on power-of-two cyclotomic rings:

$$R_q = \exists_q / \langle x^n + 1 \rangle$$

The optimized Residue Number System (RNS) variants show significant increase performance compared to their previous respective implementations [15]. The Residue Number System works with native (machine-word size) integers because it is faster than multi-precision integer arithmetic. The Residue Number System breaks rings of large bit-width integers into a parallel set of rings (<64-bit residues) allowing performant computation on 32/64-bit CPU architecture.

Modulus q is represented as product of integers:

$$q = \prod_{i=1}^{k} q_i$$

Modulus q is a functional parameter which determines how many computations are allowed without the application of bootstrapping procedure [16].

One of the most important feature of the homomorphic encryption mechanisms is that they add noise to a ciphertext during performing encryption, multiplication, addition and rotation homomorphic operations. Homomorphic operations, especially multiplication, increase level of the noise. If the noise becomes too large ciphertext can not be decrypted successfully. Noise budget is the total amount of noise that can be added until the decryption fails [17]. The bootstrapping is the procedure of "refreshing" a ciphertext by running homomorphically decryption operation that reduces level of noise.

All analysed homomorphic encryption schemes (BGV, BFV, CKKS) support the following homomorphic operations [18]:

- Addition;
- Multiplication;
- Rotation.

AN EXAMPLE OF APPLIANCE OF HOMOMORHIC ENCRYPTION

The data privacy concerns are increasingly affecting the Internet of Things (IoT) in which it is very challenging to protect the privacy of the underlying data [7]. Functional architecture of IoT platform, its core decription and an example of its application can be found in [19]. An overview of secure model of SOA based healtcare systems with mobile web service is shown in [20]. In the model data the medical data (DNA, patient data) can be protected using homomorphic encryption mechanisms.

It can be detected following classes of attacks on systems based on IoT devices and cloud servers [7]:

- Network attacks. A network attacker sees all traffic across all non-private networks. A network attacker may act actively or passively and may exploit side-channel leakage transmitted over the network to learn user data.
- Cloud attacks. A cloud attacker may refer either to an external entity who can corrupt a cloud sever or a cloud infrastructure provider themselves. A cloud attacker may attempt to bypass traditional protections in the cloud to obtain read access to private data through conventional software methods (e.g. buffer overflows), side-channel attacks such as timing or power analysis, or physical attacks.
- Device attacks. The attack refers to an attack on the device itself. These attacks may be remote and injected through direct or indirect software attacks.

Improvements and optimization of homomorphic encryption operations can make homomorphic encryption schemes enough efficient for practical usage in real systems.

Encryption is asymmetric, while evaluation refers to computation on encrypted data (performed by cloud based server) to provide encrypted results. Finally, decryption can be performed either by a trusted party (Smartphone, PC) or by same IoT device (Figure 4).

The before mentioned homomorphic encryption schemes support a technique called batching. The batching technique refers to the encoding of multiple messages into a single plaintext. The resulting batched plaintext can be encrypted into a single corresponding batched ciphertext.

Moreover, computation on batched ciphertexts can be performed in a Single Instruction Multiple Data (SIMD) way on the underlying messages, reducing the cost of homomorphic evaluation by several orders of magnitude.

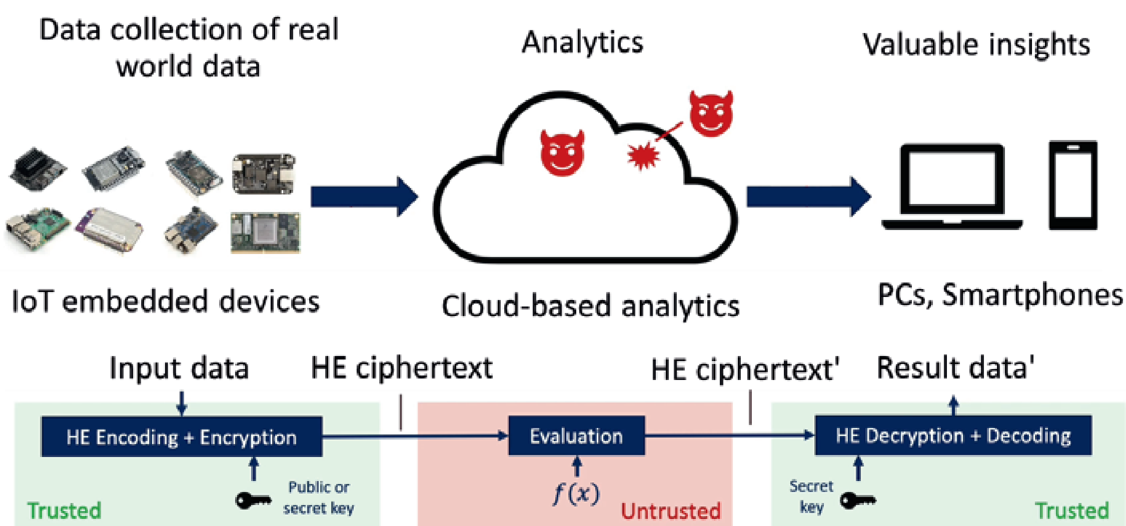


Figure 4: *Appliance of HE with IoT devices*

HOMOMORPHIC SCHEMES

The homomorphic encryption scheme named BGV was proposed in [3]. The BGV is a levelled homomorphic encryption scheme, meaning that the parameters of the scheme depend on the multiplicative depth that the scheme is capable to evaluate. Multiplicative depth of BGV scheme determines how many sequential multiplications can be performed.

The BFV scheme, presented in [4] is a homomorphic cryptographic scheme based on the Ring-LWE problem in a cryptographic lattice.

The homomorphic encryption scheme named CKKS, presented in [5] is known as Homomorphic Encryption for Arithmetic of Approximate Numbers (HEAAN). Operations supported in CKKS homomorphic encryption scheme are shown in Figure 5. The

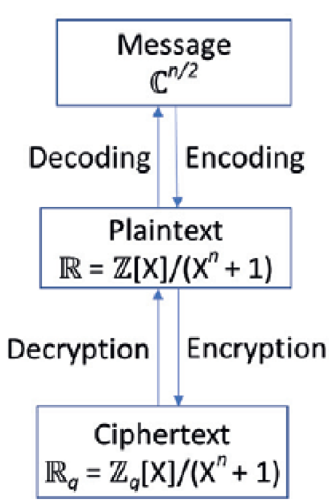


Figure 5: *Homomorphic encryption operations in CKKS*

homomorphic encryption CKKS scheme enables computations on vectors of complex values.
The approximate homomorphic encryption scheme CKKS has the following characteristics:

- $Dec(Enc(m)) \approx m$;
- $Dec(ct_1 * ct_2) \approx Dec(ct_1) * Dec(ct_2)$;
- Noise bounds are determined by the parameter set.

In the CKKS scheme noise is considered as a part of numerical error in approximate computation. It supports homomorphic rounding-off.

EXPERIMENTAL ANALYSIS

In the experimental analysis we measure the time needed for performing the following homomorphic operations: Public key encryption (Table 1), Secret key decryption (Table 2), Homomorphic addition (Table 3), Homomorphic multiplication without relinearization (Table 4) and Homomorphic rotation (Table 5). We evaluated the use of the following homomorphic mechanisms:

- BGV,
- BVF, and
- CKKS;

which are implemented in the following open-source libraries respectively:

- Microsoft SEAL [21] and
- Palisade [16];

Palisade [16] is multi-threaded library written in C++ 11. It uses the Number Theory Library (NTL) [22] to accelerate underlying mathematical operations. Palisade supports more schemes, including

BFV, BGV, CKKS. Additionally, Palisade supports multi-party extensions of certain schemes and other cryptographic primitives like Proxy Re-Encryption (PRE) and digital signatures [9].

Microsoft Simple Encrypted Arithmetic Library (SEAL) [21] is a homomorphic encryption library that allows addition and multiplication operations on encrypted integers or real numbers. Microsoft SEAL is written in C++11 and contains a .NET wrapper library for the public API. The latest available version is 4.0.0.

The homomorphic encryption code was performed on a Raspberry Pi 4 model B IoT device with:

- Broadcom BCM2711 quad-core Cortex-A72 (ARM v8) 64-bit SoC @ 1.8GHz;
- 4GB LPDDR4-3200;
- Bluetooth 5.0, Bluetooth Low Energy (BLE);
- Debian GNU/Linux 11 (bullseye);
- Hardware model: BCM2835.

All homomorphic encryption operations are performed with homomorphic encryption ciphertext dimension n = 8192.

In tables 1, 2, 3, 4, and 5 show the results of public key encryption, secret key decryption, homomorphic encryption addition, homomorphic encryption multiplication, and homomorphic encryption rotation tests respectively, where:

- Times in the last two columns (HE Library) are expressed in microsecond (μs);
- Each homomorphic encryption related operation performed 1000 times;
- We used 128-bit homomorphic encryption security level. The homomorphic encryption standard set with more than 128 bits of security with reference to classical cryptography computer attacks.
- Ciphertext dimension is n.

The public key encryption operation in all tested homomorphic encryption schemes BFV scheme has the best performance when the Palisade library is used.

Table 1: *Public key encryption in IoT device*

HE schemes	HE parameter	HE library	
	Ciphertext dimension n	Palisade	SEAL
BFV	8,192	7,820 μs	14,250 μs
BGV	8,192	8,511 μs	14,189 μs
CKKS	8,192	8,209 μs	16,888 μs

In Raspberry Pi 4 IoT device the public key encryption (Table 1) is achieved faster performance using Palisade library than SEAL library.. In Palisade the public key encryption the best performance is achieved using BFV homomorphic encryption scheme. In SEAL the public key encryption the best performance is achieved using BGV homomorphic encryption scheme.

Table 2: *Secret key decryption in IoT device*

HE schemes	HE parameter	HE library	
	Ciphertext dimension n	Palisade	SEAL
BFV	8,192	1,607 μs	5,285 μs
BGV	8,192	2,094 μs	5,089 μs
CKKS	8,192	11,644 μs	889 μs

The secret key decryption (Table 2) in CKKS scheme has better performance in the SEAL than in the Palisade library.

The secret key decryption of exact homomorphic encryption schemes (BGV, BFV) faster performance is achieved in Palisade than in SEAL library. In Palisade library secret key decryption is fastest by using BFV homomorphic encryption scheme and in SEAL the operation is fastest by using CKKS approcimate homomorhic encryption scheme.

Table 3: *Homomophic encryption addition in IoT device*

HE schemes	HE parameter	HE library	
	Ciphertext dimension n	Palisade	SEAL
BFV	8,192	358 μs	1,280 μs
BGV	8,192	493 μs	1,341 μs
CKKS	8,192	808 μs	1,632 μs

In Raspberry Pi 4 IoT device the homomophic encryption addition (Table 3) is achieved faster performance using Palisade library than SEAL library. In both library the best performance of homomophic encryption addition is achieved using BFV scheme.

Table 4: *Homomophic encryption multiplication in IoT device*

HE schemes	HE parameter	HE library	
	Ciphertext dimension n	Palisade	SEAL
BFV	8,192	17,845 μs	57,332 μs
BGV	8,192	1,317 μs	16,605 μs
CKKS	8,192	1,447 μs	3,293 μs

In Raspberry Pi 4 IoT device the homomophic encryption multiplication (Table 4) is achieved faster performance using Palisade library than SEAL library. In Palisade the homomophic encryption multiplication is achieved the best performanse using BGV scheme. In SEAL the homomophic encryption multiplication is achieved the best performanse using CKKS scheme.

Table 5: *Homomophic encryption rotation in IoT device*

HE schemes	HE parameter	HE library	
	Ciphertext dimension n	Palisade	SEAL
BFV	8,192	3,940	16,207
BGV	8,192	3,641	17,482
CKKS	8,192	9,307	16,288

In Raspberry Pi 4 IoT device the homomophic encryption rotation (Table 5) is achieved faster performance using Palisade library than SEAL library. In Palisade the homomophic encryption rotation is achieved the best performanse using BGV scheme. In SEAL the homomophic encryption rotation is achieved the best performanse using BFV scheme that is slightly faster than in case of appliance the operation using CKKS.

Conclusions

Homomorphic encryption enables performing computations on the encrypted data, without decrypting them. The work compares the time needed to execute homomorphic operations, like, public key encryption, secret key decryption, addition, multiplication, and rotation implemented in the open-source libraries: Microsoft SEAL and Palisade. The operations are compared for BGV, BFV and CKKS homomorphic encryption schemes implemented in the libraries.

Homomorphic operations that are usually performed at client side: public key encryption and secret key decryption in case of appliance exact arithmetic homomorphic operations are achieved best performance using Palisade library. In case of usage of approximate arithmetic operations (CKKS) public key encryption operations have better performance in Palisade than in SEAL library. The SEAL library provides better performance in secret key decryption operation.

To demonstrate the possibility of implementation of homomorphic encryption related operations in IoT device we deployed the different homomorphic encryption schemes on Raspberry Pi 4 model B based IoT platform. Our results show that homomorphic encryption operations can be applied on embedded devices.

The CKKS homomorphic encryption scheme can be considered as natural choice for IoT devices since it can efficiently perform secure computation on the type of real-valued data often sampled by sensors. In this sense in future work, it can be considered method of efficient implementation of CKKS scheme especially operations of public key encryption and private key description that are performed on client (IoT) side.

The performance of current fully homomorphic encryption schemes, especially for large parameters, can still be improved. Further improvement can be achieved by implementation low-level homomorphic operations in an assembly language which is executed on Raspberry Pi 4 IoT hardware platform.

References

[1] A. Acar, H. Aksu, A. Selcuk, and M. Conti, "A Survey on Homomorphic Encryption Schemes: Theory and Implementation," ACM Comput. Surv. 1, 1, Article 1, http://dx.doi.org/10.1145/3214303, 2018.

[2] V. Lyubashevsky, C. Peikert, and O. Regev, "On ideal lattices and learning with errors over rings," Journal of the ACM (JACM) 60, no. 6, 2013.

[3] Z. Brakerski, C. Gentry, and V. Vaikuntanathan, "Fully Homomorphic Encryption without Bootstrapping," Cryptology ePrint Archive, Report 2011/277. https://eprint.iacr.org/2011/277, 2011.

[4] J. Fan and F. Vercauteren, "Somewhat practical fully homomorphic encryption," IACR Cryptology ePrint Archive, 2012:144, 2012.

[5] J.H. Cheon, A. Kim, M. Kim, and Y. Song, „Homomorphic encryption for arithmetic of approximate numbers," Cryptology ePrint Archive, Report 2016/421, https://eprint.iacr.org/2016/421, 2016.

[6] W. Yuan, H. Gao, "An Efficient BGV-type Encryption Scheme for IoT Systems," https://doi.org/10.3390/app10175732, 2020.

[7] D. Natarajan, W.Dai, "SEAL-Embedded: A Homomorphic Encryption Library for the Internet of Things", IACR Transactions on Cryptographic Hardware and Embedded Systems, Vol. 2021, pp 756-779.

[8] H.Seo, H.Kwon, Y.Kwon, K.Kim, S.Choi,H.Kim, K.Jang, "Fast Number Theoretic Transform for Ring-LWE on 8-bit AVR Embedded Processor," Sensors (Basel). 2020 Apr 5;20(7):2039. doi: 10.3390/s20072039.

[9] G. Đorđević, M. Marković, P.Vuletić, "Performance comparison of homomorphic encryption scheme implementations", 8th International Conference on Electrical, Electronic and Computing Engineering, IcETRAN 2021, pp 514-519.

[10] C. Aguilar Melchor, M. Kilijian, C. Lefebvre, T. Ricosset, "A Comparison of the Homomorphic Encryption Libraries HElib, SEAL and FV-NFLlib," in: Lanet JL., Toma C. (eds)

Innovative Security Solutions for Information Technology and Communications. SECITC 2018. Lecture Notes in Computer Science, vol 11359. Springer, Cham. https://doi.org/10.1007/978-3-030-12942-2_32, 2019.

[11] T. Maha, S. Hajji, and A. Ghazi, "Homomorphic encryption applied to the cloud computing security," in Proceedings of the World Congress Engineering, vol. 1, pp. 4-6, 2012.

[12] L. Ducas and D. Micciancio, "FHEW: bootstrapping homomorphic encryption in less than a second," in E. Oswald and M. Fischlin, editors, Advances in Cryptology – EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, So_a, Bulgaria, April 26-30, 2015, Proceedings, Part I, volume 9056 of Lecture Notes in Computer Science, pages 617-640. Springer, 2015.

[13] I. Chillotti, N. Gama, M. Georgieva, and M. Izabachene, "Faster packed homomorphic operations and e_cient circuit bootstrapping for tfhe," in Advances in Cryptology-ASIACRYPT 2017: 23rd International Conference on the Theory and Application of Cryptology and Information Security, pages 377-408. Springer, 2017.

[14] O. Regev, "The learning with errors problem," in Blavatnik School of Computer Science, Tel Aviv University Invited survey in CCC, 2010.

[15] J. H. Cheon, K. Han, A. Kim, M. Kim, and Y. Song, "A full rns variant of approximate homomorphic encryption," Cryptology ePrint Archive,Report 2018/931, https://eprint.iacr.org/2018/931, 2018.

[16] Y. Polyakov, K. Rohloff, G.W. Ryan, and D. Cousins, "PALISADE Lattice Cryptography Library User Manual (v1.10.6)", 2020.

[17] S. Sathya, P. Vepakomma, R. Raskar, R. Ramachandra, and S. Bhattacharya, "A Review of Homomorphic Encryption Libraries for Secure Computation," http://arxiv.org/abs/1812.024, 2018.

[18] M. Albrecht, M. Chase, H. Chen and others, "Homomorphic encryption standardization," homomorphicencryption.org, 2018.

[19] G. Đukanović, G. Popović, D. Kanellopoulos, "Scaling complexity comparison of an ACO-based routing algorithm used as an IoT network core," Journal of Information Technology and Applications, JITA 10(2020) 2:73-80, (UDC: 004.738.5:004.738.057.4), (DOI: 10.7251/JIT2002073DJ), Volume 10, Number 2, Banja Luka, December 2020 (69-128), ISSN 2232-9625 (print), ISSN 2233-0194 (online),

[20] G. Đorđević, M. Marković, "On Possible Cryptographic Optimization of Mobile Healthcare Application," JITA – Journal of InformationTechnology and Applications Banja Luka, JITA 9(2019) 2:80-88, (UDC: 004.056.55:621.39), (DOI: 10.7251/JIT1902080DJ), Volume 9, Number 2, Banja Luka, december 2019.(49-128), ISSN 2232-9625 (print), ISSN 2233-0194 (online)

[21] K. Laine, "Simple Encrypted Arithmetic Library 2.3.1," 2017.

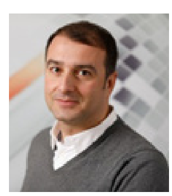
[22] V. Shoup and others, "NTL: A library for doing number theory," http://www.shoup.net/ntl.

Received: January 17, 2022.
Accepted: May 24, 2022.

About the authors

Goran V. Đorđević was born 1972 in Novi Sad, Serbia. He received a BSc in Computer Science at the Technical Military Academy in 1996. Afterwards he did his post-graduate studies, at the Faculty of Electrical Engineering of University of Belgrade where he received a MSc. Currently employed as a senior software developer in AET Europe. His main areas of interest are smart card security and smart card applications, security protocol design, mobile devices, tokens, Internet of Things and information security.

Milan Marković received B.S.E.E., M.S.E.E., and Ph.D. degrees in electrical engineering from Faculty of Electrical Engineering, University of Belgrade, Serbia, in 1989, 1992, and 2001, respectively. He is an Associate Professor on College of Information Technology, Pan-European University of Apeiron in domain of information security courses. His research interests are mainly in public key infrastructure, information security, cryptographic algorithms, mobile security, identity management, secure e/m-banking and e/m-government, trust services, ISMS, Blockchain, etc. He has published more than 320 scientific papers.



Pavle Vuletić obtained his BSc, MSc and PhD in Computer Systems and Network Architecture from University of Belgrade, School of Electrical Engineering (ETF). He used to work on all positions from network engineer to the deputy director of AMRES, national research and education network where he participated in the establishment of the first national CSIRT team. He is currently the associate professor at the ETF at the Department of Computer Engineering and Information Theory, teaching Data Security, Computer Systems and Network Security, Advanced Computer Networks and Software Defined Networks courses and the head of the Laboratory for Information Security at the ETF. His research interests span from network and systems security, intrusion detection, network and system performance and monitoring to software defined networks and network and systems management.

For citation

Goran Đorđević, Milan Marković, Pavle Vuletić, Evaulation of Homomorphic Encryption Implementation on Iot Device, *JITA – Journal of Information Technology and Applications, Banja Luka*, Pan-Europien University APEIRON, Banja Luka, Republika Srpska, Bosna i Hercegovina, JITA 12(2022) 1:32-39, (UDC: 004.738.5:621.391), (DOI: 10.7251/JIT2201032DJ), Volume 12, Number 1, Banja Luka, June (1-64), ISSN 2232-9625 (print), ISSN 2233-0194 (online), UDC 004

Improving the Process of Online Education by Introducing Innovation

Boris Kovačić

Pharmaceutical Chamber of the Republic of Srpska, kovacic3boris@gmail.com

Branko Latinović

Pan-European University APEIRON, branko.b.latinovic@apeiron-edu.eu

Critical Review

https://doi.org/10.7251/JIT2201040K UDC: 004.738.5:37.091.3

Abstract: The aim of this paper is to point out the information chain of supply to graduate pharmacists and masters of pharmacy, members of the Pharmaceutical Chamber of the Republic of Srpska. In addition, point to CRM in order to achieve greater satisfaction of members. Point out solutions and innovations that improve data exchange of associations that provide continuous education of graduate pharmacists and masters of pharmacy, members of the Pharmaceutical Chamber of Republika Srpska, as well as online platforms for online education of the Pharmaceutical Chamber of Republika Srpska and local databases.

Keywords: Information supply chain, CRM, Online education, Moodle, data exchange, data exchange difficulties.

Introduction

The Pharmaceutical Chamber of the Republic of Srpska is an independent, professional organization of masters in pharmacy, graduated pharmacists and graduated biochemists. As an independent professional organization, it first appears on the day of its founding. Previously, before gaining independence status, it was part of a unique Health Chamber. The Medical Chamber consisted of: doctors of medicine, dentistry doctors and graduated pharmacists. Within the Health Chamber were three councils that made up the basic structure of that chamber, with 21 elected representatives, each from the ranks of medical doctors, dentistry doctors and graduated pharmacists. Other bodies were selected in accordance with the legal regulations and the Statute and the Health Chamber. [4]

On May 17, 2003, the National Assembly of Republika Srpska adopted the Law on Health Chambers, which legally established the Pharmaceutical Chamber of Republika Srpska. [6]

At the assembly meeting held at the Medical Electronics offices in Banja Luka, the Pharmaceutical Chamber of Republika Srpska was established on July 9, 2003. On this occasion, bodies of the Chamber were elected: the Assembly, the Executive Board, the Supervisory Board and the Court.

The Pharmaceutical Chamber of Republika Srpska has a fundamental role to ~~do to~~ continuously educate pharmacists in Republika Srpska. Members of the Pharmaceutical Chamber of Republika Srpska are masters of pharmaceuticals, graduated pharmacists and graduated biochemists, and other health professionals. Members of the Pharmaceutical Chamber of Republika Srpska attend symposiums, vocational lectures, and thus collect certain points. After a certain period, when a member collects enough points, the Pharmaceutical Chamber of Republika Srpska issues a member a license to work for a period of one to five years. The specified license serves as a proof of the ability of a member, a healthcare professional, who has attended professional continuous educations for health professionals for a period of time, and meets the requirement that he can continue to work as a healthcare professional.

Education was carried out by the Pharmaceutical Chamber of Republika Srpska through lectures at professional meetings, organizing symposiums and other forms of face-to-face education. In addition, the Pharmaceutical Chamber of Republika Srpska issued a printed yearbook of works.

The website was initially created in a html code and contained basic information about the Republika Srpska Chamber of Pharmacy.

ONLINE EDUCATION PROCESS

The initial project was done for just over half a year. When starting the initial part of the project, the request was quite confusing in terms of the layout and design of the website. The initial project was supposed to allow the transition from the HTML platform to php/mysql, with the CMS platform Joomla version 1.5 used in the specific solution. According to the fact that the production of the template for the Joomla platform was quite complex, and in the circumstances that it was not a definite strictly task, 3 solutions of the site design were made and the selected version, which was accepted, continued to be refined for another 3 months.

In the test period of one year, the news and database of pharmacies and wholesalers in Republika Srpska were updated and placed on the site electronic forms in doc and pdf format.

Forms that are electronically prepared allow you to fill in the requirements for issuing or renewing a license or changing the workplace. Within the forms there is a list of the necessary documentation that needs to be attached for the purpose of issuing a license or other document. These electronic forms in PDF format can be printed and then manually filled in, or filled electronically using the program “Acrobat Reader” which is a free program.

After registration in the register of members of the Pharmaceutical Chamber of Republika Srpska, an user profile for access to “online” education is opened.

Materials from professional lectures and Symposiums, e-books and publications, The Proceedings of Professional Papers have been entered.

No functionality of quizzes, forums, or other functionality is required when building, but only the above educational functionality, although the system has these functionality.

Lexi-Comp base has been incorporated into the system. The English-language database contains dosing and a way of issuing a particular drug given the patient’s age and weight. Besides, this base provides information about the interaction of food and medicines, and medicine with medicines. The database is updated at any moment and provides the latest data. Data update, in English, is run by several thousand pharmacists worldwide and about 15,000 clinical pharmacologists, who update it daily. “Lexi-Comp” base in Serbian is much more modest and contained an incorporated registry of medicines of Serbia. The reason for writing in the past tense is that the license has expired, and according to the last available information that this version no longer exists.

Educations placed from multiple symposiums are in pdf format, not only a collection of papers, but also presentations of lecturers that are also set in pdf format, to protect the rights of authors. Updating “online” education is done during the Symposium or other professional meeting, or after the meeting, depending on the need for data processing. Namely, certain lecturers submit a presentation with video content within it, and require compression or separation from the presentation with a link to video content, for the possibility of display in the “online” system.

PLANNING AND DATA COLLECTION

In order to make an improvement, it is necessary to collect the data, process it and then plan it. Information supply chain and CRM system. The first step in collecting data is the basic request of the Expert Team of the Pharmaceutical Chamber of Republika Srpska. After that, during planning, periodic meetings are additionally held to better interact and resolve segments of the system for the purpose of better planning.

REQUIREMENTS

The first and basic request was the exchange of data between the local database of the Pharmaceutical Chamber of Republika Srpska and the Online Education of the Pharmaceutical Chamber of Republika Srpska. The exchange type would be done in a way that would be basic information such as personal data, with additional data such as license number, license expiration date, number of points and finances.

The second request was the exchange of data between the local database of the Pharmaceutical Chamber of Republika Srpska and the Online Education of the Pharmaceutical Chamber of Republika Srpska with the expansion of functional display with the list of education of members of the Chamber. This type of education would be done as before, with an additional data series. Namely, this edited data set would be connected to each member individually. A data series would have to be tied to a separate table with a relationship of 1 to more, or more educations could have been had by each member. The education list is accompanied by data on the number of points, as well as other data that tracks each education individually.

The third request was the exchange of data between the local database of the Pharmaceutical Chamber of Republika Srpska and the Online Education of the Pharmaceutical Chamber of Republika Srpska and the exchange of education data of other professional associations, legal entities and higher education institutions that have the authority to hold this type of education.

Meetings

The meetings allowed for a better understanding of the topic, and the breakdown of data into finer parts and better description of metadata. Periodic meetings were held for the purpose of controlling the flow of the project, and as a form of fragments merged into a whole.

CRM

The term CRM comes from the English abbreviation "*customer relationship management*". So it arises that the CRM system is a system of data management with service users. The CRM system enables legal entities to collect, analyze, edit, and store information about potential and current users.

Art and science that ensures the right product or service enters the hands of the right customer in the right quantity and at the right time is known as supply chain management (SCM). [2]

The three elements of CRM success are:

- people (employed by the company);
- processes (reengineering of the process must be carried out in the company);
- technology (the company must choose the correct technology to conduct the process). [1]

This work discusses the MODEL CRM B2C, respectively dedicated user.

In order to gather as much data as possible when developing the system, a survey was conducted, then a system improvement was planned by adding chats and quizzes. Namely, the chat will enable direct interaction of the Pharmaceutical Chamber of Republika Srpska with users. Additional information can be obtained through short quizzes. The separate system of online education of the Pharmaceutical Chamber of Republika Srpska, which is of closed type, i.e. only for members of the Pharmaceutical Chamber of Republika Srpska, has additional functionality that the Moodle CMS system offers, which is to monitor interaction through clicks and visits of certain content. This enables maximum interaction and customer satisfaction.

By integrating all modules, processes, links into a single system, we get a high-level CRM system.

Poll

A sample of about 1,000 members was used to create this survey. At the time of this survey, which was taken from the master's thesis, only about 10% of the observed sample were users of the information system service. The rationale for such a low rate of use is politics and the general situation in the country, which is why there is a lack of interest and the need only for the necessary activities. The response to the aforementioned survey was some-

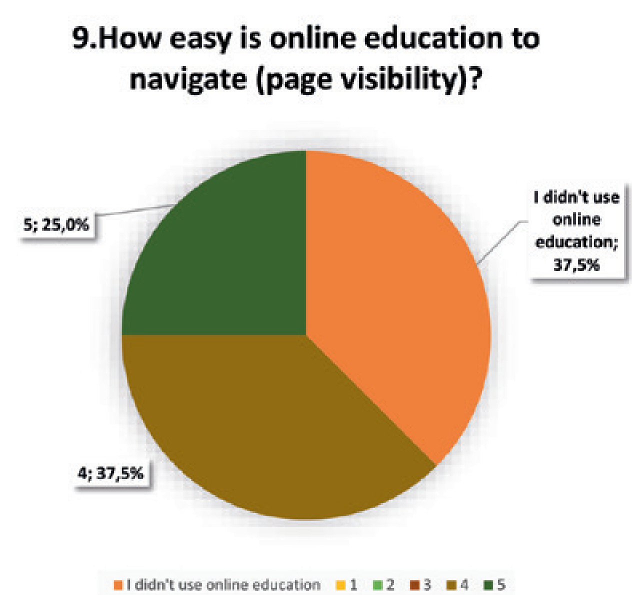


Figure: 1 *Author: Kovacic B, Diagram of simplicity of dealing with online education according to survey participants [3]* where around 2% of users. From the above, the data processed are presented graphically in the form of diagrams in further work.

Only part of the survey is presented in this work.

After basic subject data, I singled out the part related to "*Online education*" of the Pharmaceutical Chamber of Republika Srpska.

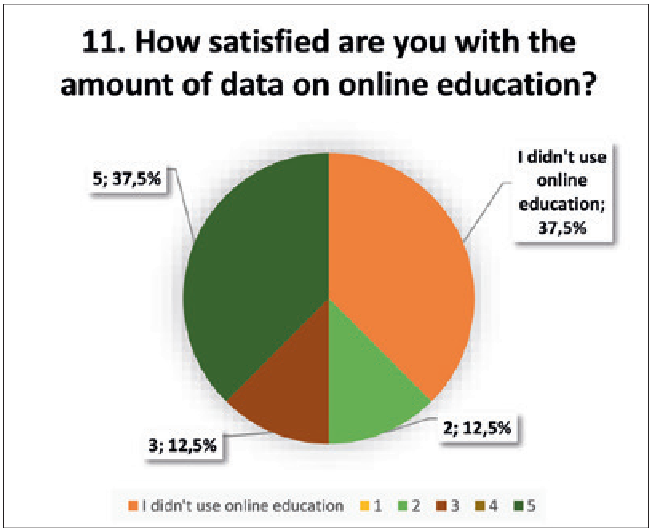


Figure: 2 *Author: Kovacic B, Satisfaction Diagram by amount of data on online education according to survey participants [3]*

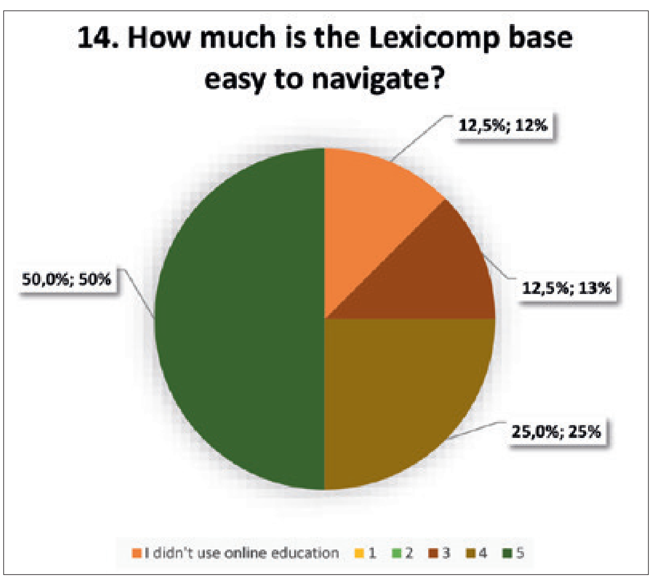


Figure: 3 *Author: Kovacic B, Diagram of Simplicity in Lexi Comp database according to survey participants [3]*

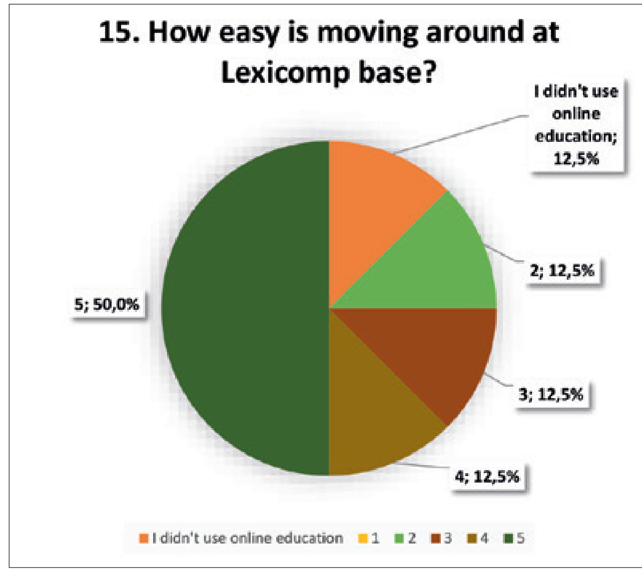


Figure: 4 *Author: Kovacic B, Diagram of ease of movement in Lexi Comp database according to survey participants [3]*

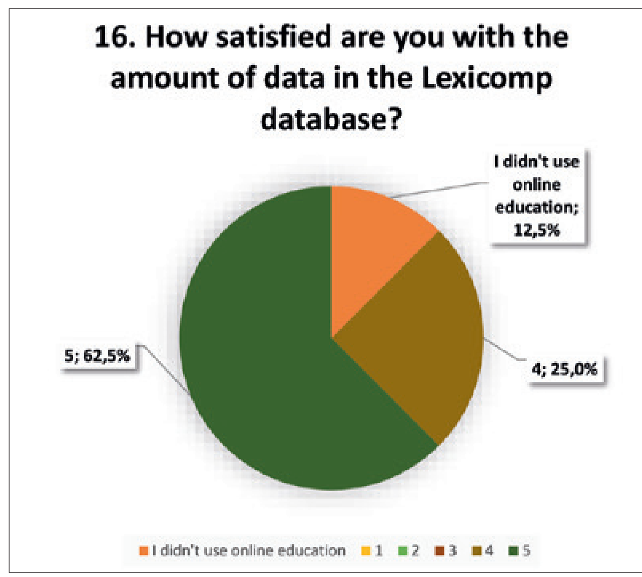


Figure: 5 *Author: Kovacic B, Satisfaction Diagram by Amount of Data in Lexicomp Database according to survey participants [3]*

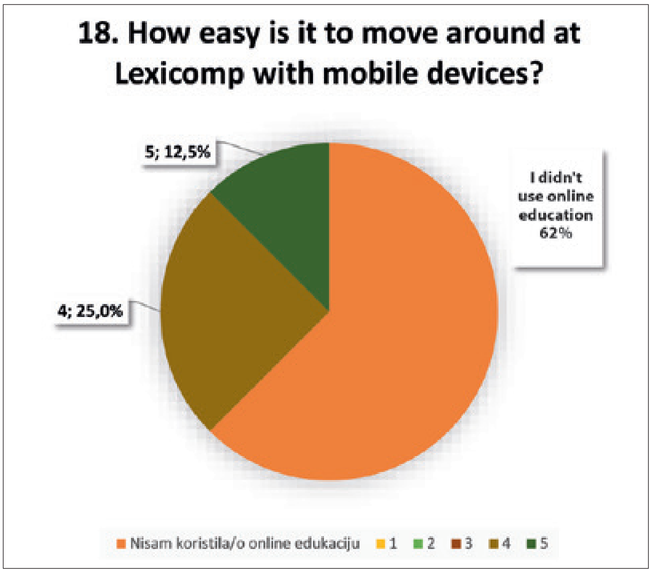


Figure: 6 *Author: Kovacic B, Diagram of ease of movement in Lexicomp database with mobile phone according to survey participants [3]*

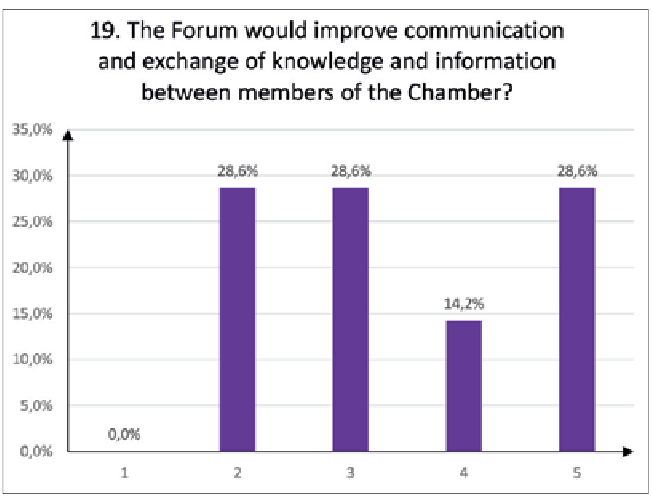


Figure: 7 *Author: Kovacic B, Diagram of improving communication and exchange of knowledge and information through forum in the opinion of survey participants [3]*

By reviewing the work, we come to the conclusion that although the path to the introduction of "online" education and its use is quite complex, it yields positive results.

When introducing "online" education, it is important to have management support, necessary resources and adequate external resources (adequate legal regulation and other parameters of society), and to use proven philosophies, adequate models and methods. Kaizen philosophy and Agile approach to software development should be used in development and planning for success. Toyota

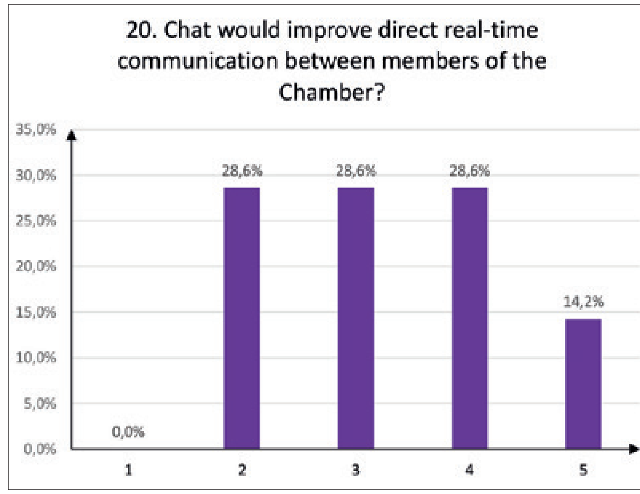


Figure: 8 *Author: Kovacic B, Diagram of communication improvements between chamber members through the introduction of chat in the opinion of survey participants [3]*

succeeded by using the Kaizen philosophy in its development. Many projects succeed, and businesses raise their quality and quantity with the help of this philosophy. The ISO 9001 stadard defines continuous improvement of business processes. With step-by-step improvement in the agile approach to software development if we use the Kaizen philosophy, which defines the continuous elimination of losses, success is almost guaranteed. In e-education it is necessary to apply the above mentioned methods and philosophies, e-education requires continuous improvement.

The insight into the survey concludes that the lifespan of this CMS system was at an end, and needed its reconstruction, which is currently under way.

Chat

In order to enable better interaction with users of the System of The Pharmaceutical Chamber of Republika Srpska, chat needs to be integrated. With the Moodle platform in use, it is possible to integrate multiple software solutions. I'll list some versions of chat that can be integrated into the Moodle platform, which are: Skype, We Chat, Microsoft Live Messenger, OpenMeetings, MyLiveChat, Global Chat, and other.

Chat, depending on the version, enables you to interact with users continuously or during the educational period. Depending on the solution, there are additional video communication restriction options.

Quiz

There is an application part in the CMS system on the Moodle platform to create quizzes. The quiz not only has to play a role in testing knowledge after listening to a course or other form of online education, but may serve as a similar survey for users' interaction with the service provider. Namely, a quiz as well as a survey can serve as a feedback, a response and thus collect information.

Supply Chain

When we mention the supply chain, we come across one of the definitions that suggests that the supply chain also includes all transfers of services and physical goods needed to produce and valorize goods on the market, i.e. to be provided, delivered to the end user.

In this work, goods are information and educational materials. Distribution of information and educational materials is done through printed materials, books, certificates, official acts, internet portals, educational pages, SMS services, e-mail, fax and other forms of communication. The focus of this work is on the timely supply of information by the Pharmaceutical Chamber of Republika Srpska. The information supply chain enables timely delivery of the information delivery service.

One form of information supply chain of the Pharmaceutical Chamber of Republika Srpska is the SMS system.

The SMS system offers an automated response, without the participation of the human factor, on the requested information of a member of the Pharmaceutical Chamber of Republika Srpska. The information requested in the query is data about the state of the points, the expiration of the license and the finances related to the cost of the membership fee, which are delivered in real time. This type of automatic SMS system response is because it is connected to the database. Query time is not limited, so at any time it enables a query about the requested data and a real-time response about the state of the database data for the specified member. Instructions have been made for using sms systems. The above instruction was presented in several professional lectures, as well as at the Symposiums of Pharmacists. The instruction explains, step by step, how to use the system, and what is the number to which the SMS query is being sent. In this way, all members of the Pharmaceutical Chamber of Republika Srpska, as users of sms system, are aware of the use and functionality of the specified system.

In addition to the base, the SMS system is connected to the application, so that the Professional Service of the Republika Srpska Chamber of Pharmacy can inform its members in a particular region about professional lectures in the region. [3]

Promotion and innovation

Under the notion of innovation, we will find the definition that innovation is the use of something new or significantly improved, a particular good or service. In this work, the point of innovation is to exchange data at a higher level. For the notion of promotion, we have synonymous with progress or improvement. In this work, this term is used in terms of improving the existing system.

Initial solution:

The initial solution of "online" education of the Pharmaceutical Chamber of Republika Srpska is that the local application is related to the local base MSSQL, and "online" education on the Moodle platform is a separate system. Data synchronization is performed semi-automaticly through a daily report that is integrated manually through "Chron job" activity on the provider's server.

Innovation would mirror the rise of the local "web" server, by installing a Moodle platform with an MSSQL base in the background, and by connecting the local base and online education base, so that the enrollment process would be fully automated and integrated with the local base (application), SMS system.

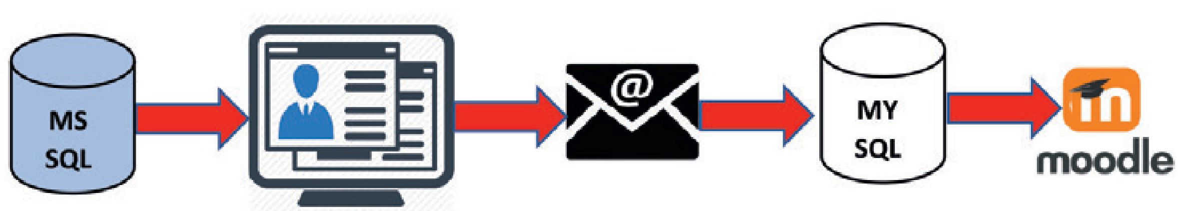


***Figure: 9** Chain of information supplied by members of The Pharmaceutical Chamber of Republika Srpska*

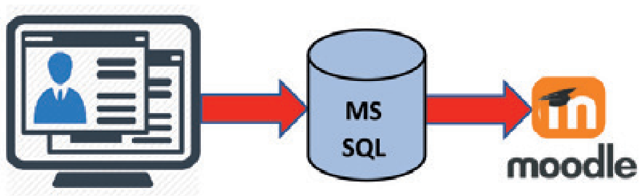


***Figure: 10** Improved chain of information supply of members of The Pharmaceutical Chamber of Republika Srpska*

The Moodle system enables import and export data in SQL, CSV, and TXT formats. The purpose of this export and import is to exchange data with other systems. In format CSV, it is possible to determine the partial used in the export document.

A CSV report is generated from the MSSQL database and the local application is sent daily via e-mail, then manually uploaded to Moodle data.

Future solution:

Phase 1

Data is typed into the same MSSQL database through a local application and automatically displayed in Moodle data because it is the same database. This type of improvement is the automation of the information exchange process.

Phase 2

However, there is a need for a step forward. Namely, pursuant to the role of the Pharmaceutical Chamber of Republika Srpska to support education by other legal entities that have the approval for education, there is an additional need for data exchange.

To achieve the final version of the supply chain, you need to even out the data exchange parameter, or key.

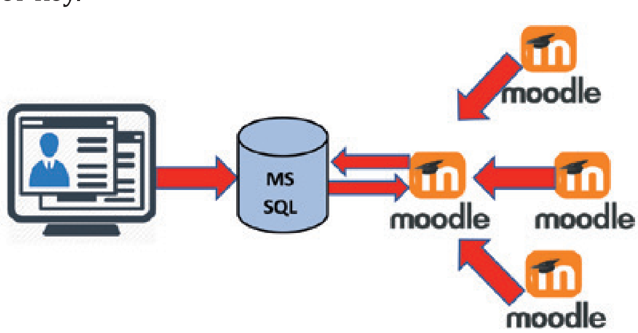


***Figure: 11** Final version of the information supply chain of members of the Pharmaceutical Chamber of Republika Srpska*

	ID	LicenseNO
1	489	01-458/II

***Figure: 12** Excerpt from local database Pharmaceutical Chamber of Republika Srpska [5]*

Pursuant to the sending of ID from the local base would not make sense, nor the ID that is in the Moodle system of the Pharmaceutical Chamber of Republika Srpska, a unified solution had to be found, which was described in one of the previous works.

If we were to review such databases at the national level of Republika Srpska, or if we were to introduce a registry of all health professionals at the Bosnia and Herzegovina level, it would be necessary to restore the forms as at the beginning, when there was a Health Chamber in Republika Srpska. 01- was the mark for a doctor of medicine 02 - was the mark for graduate pharmacists or masters of pharmacy.

The solution is to normalize the license number, for example, in form 01-1111/2 in such a way that it is in the form of a 011111 and thus becomes a pharmacist's ID in the form that is the best solution for using the key in the database, and will not be repeated where all conditions are met. [5]

Adjusting license numbers with normalization in the example shown would be converted from a 01-458/II format to a 010458 format.

	PharmacistID	LicenseNO
1	010458	01-458/II

***Figure: 13** View the future named Column PharmacistID in a local database [5]*

A step further in the up-and-coming way allows me to stop Moodle by traveling his add-ons and sprayed syringes. The new mode of the Moodle syringe, in addition to offering an app for the cotton platform, also has syringes such as the provinciality of the validity of the certificate. Provincial is performed in 2 ways. One of the ways in which QR is loaded is to the website of the ma'am confirming the validity of the certificates. Another way is to load the website by figuring out the validity of the certificates by typed in a short code, which is given to maintained algorithms by random choice. This short code is visible in an image labeled 9DwSV9F2PD.

***Figure: 14** Snippet of certificate appearance through advanced services (document is issued in Serbian)*

Conclusion

Modern technologies in the pure supply chain of information enable real-time information supply. It is also important to note that the collaboration of data between the systems and their compatibility is important for data exchange, and makes it easier for the user to view the data in one place, at one point. CRM is important in response to the success of reingnostication and innovation in the system. Innovation in this work refers to the collaboration of data between systems to achieve a higher level of CRM and better information supply. Finally, let me emphasize that besides technologies, the key link in the human resources system are human resources.

Literature

[1] Latinovic Branko, ERP – materials from lectures, Pan-European University Apeiron, Faculty of Information Technologies;

[2] Marianne Bradford, Ph.D, Professor Poole College of Management North Carolina State University Raleigh, NC Copyright 2015, Modern ERP;

[3] Kovacic B, Master's Work on the Subject Introduction of information communication technologies in companies to achieve quantity and raise quality, UDK 007:005.6(043.2), 25.12.2017, Pan-European University Apeiron, Faculty of Business Economics;

[4] Kovacic B., Smailovic N., Professional gathering of Information Technology for e-education, XI international scientific conference ITeO, Paneuropean university Apeiron, (proceedings), page 158-166;

[5] Avramovic Z. Zoran, Kovacic A. Boris, Multimedia database for continuous education of graduate pharmacists in the Republic of Srpska, conference proceedings pages: 39-43 Complex interaction of linguistic and graduating departments in a technical university Moscow, May 27, 2021 Organizers: Russian University of Transport (MIIT), UDK: 378: 371.4;

[6] https://fkrs.org/?p=645, (visited on June 3).

Received: January 28, 2022
Accepted: April 28, 2022

About the authors

Boris Kovačić was born in Bihac in 1979. He finished elementary and technical school in Banja Luka. He graduated at the Faculty of Information Technology and the Faculty of Business Economics at Pan-European Apeiron University. He has a master's degree in economics from Pan-European Apeiron University. He is employed in The Agency for Medicinal product and Medical Devices of Bosnia And Herzegovina as IT consultant.

Branko Latinović was born on April 28, 1956, in Prijedor, Bosnia and Herzegovina. He graduated from the Faculty of Economics in Banja Luka in 1980. At the same faculty he enrolls a master's degree that ends in 1994, and in 1997 successfully defended his doctoral dissertation. He works as a dean of the Faculty of Information Technologies of Pan-European University "Apeiron" in Banja Luka.

For citation

Boris Kovačić, Branko Latinović, Improving the Process of Online Education by Introducing Innovation, *JITA – Journal of Information Technology and Applications, Banja Luka*, Pan-Europien University APEIRON, Banja Luka, Republika Srpska, Bosna i Hercegovina, JITA 12(2022) 1:40-47, (UDC: 004.738.5:37.091.3), (DOI: 10.7251/JIT2201040K), Volume 12, Number 1, Banja Luka, June (1-64), ISSN 2232-9625 (print), ISSN 2233-0194 (online), UDC 004

IoT – Company Approach to IoT Modeling and Applications

Dražen Marinković, Zoran Ž. Avramović

Pan-European University APEIRON, Banja Luka, Republika Srpska, BiH

{drazen.m.marinkovic, zoran.z.avramovic}@apeiron-edu.eu

Critical Review

https://doi.org/10.7251/JIT2201048M UDC: 621.391:004.738.5

Abstract: Using the available literature, this paper attempts to present the company 's approach to IoT modeling and the incorporation of IoT technologies into business processes. Furthermore, a comprehensive overview of IoT technologies and systems of large corporations (Yokogawa, Intel) and commercial access to IoT technologies is provided. In conclusion, based on previous knowledge and scientifically based arguments, the advantages and disadvantages of IoT technologies are presented.

Keywords: IoT, IioT, Digital Intelligence, Total Cost of Ownership, Operations Excellence, Cloud, CloudIoT.

Introduction

Internet of things (or Internet of Intelligent Devices) is a term that today refers to billions of physical devices around the world equipped with sensors and software, which are connected to the Internet in order to collect and share data and information obtained. In this way, the devices enable the so-called "digital intelligence", i.e. devices are given the opportunity to use data at appropriate times without human participation, but also with the help of these data, these devices can independently regulate their work. In short, IoT is a global network that connects smart things. All IoT devices, which are interconnected, participate in communication. Most definitions of IoT technology refer to the combination of everyday things that a person uses on the Internet for the purpose of measuring, collecting, storing and exchanging data with other things and people, that is, the activity of IoT devices is based on collected data.

Distribution Of IoT Technologies

IoT technologies represent a kind of turning point in the future of computer technologies and communications, and the development of these technologies depends on technical innovations in many important areas, from wireless to nano technology. IoT devices are becoming more and more prevalent, so today we have more IoT devices than the number of people in the world. Analyst company Statista announced that by 2025, the number of IoT devices will be around 38.6 billion. According to a report by analytics firm Gartner, more than half of IoT devices are consumer products (smart TVs, smart speakers and other home devices, and the best-selling are smart electric meters and commercial security cameras).

The extent to which IoT technology is widespread and implemented in everyday activities, and the importance of IoT technology for business, is shown by the fact that since 2011 IoT week has been held on a global level. The first conference was held in Barcelona in 2011. IoT makes our homes, offices and our vehicles smarter, more measurable and easier to use. Autonomous vehicles and smart cities, which could be created thanks to IoT technology, could change the way we build or manage public space in the future. On the other hand, what is considered a shortcoming of this system is the fact that many of these innovations intrude too much on our privacy and can have too much of an impact on our private lives. The Internet of Things system consists of smart devices that the user manages through an application through the Internet. The user does not have to communicate directly with things, but with a service that stores and processes data obtained from smart devices and manages the obtained data. IoT consists of any device that is connected to the Internet. IoT enters all spheres of our lives, and gives us the opportunity to improve the quality of life, products and services, but also reduces maintenance costs, increases the efficiency of business processes, creates new types of jobs, while, on the other hand, there is the possibility of losing jobs, as there is not a man, but a machine at the heart of IoT technology. Also, one of the disadvantages of the IoT is the loss of control over a large amount of data and information, and the misuse of a large amount of private data can occur. Despite these shortcomings, IoT technology is of great interest to creators and users, but also to companies that, with a combination of cameras, sensors and connected devices, receive information about our habits, movements, emotions, regardless of whether we want it or not.

For each of us, who has a smartphone that has GPS, the location of the device is known at all times, as well as its user. Due to all the above, the question arises of monitoring and controlling all this information that the IoT sends about us. It is assumed that it will not be possible to control all the collection of information, so IoT legislation is also required.

Yokogawa Approach

Major vendors of automated equipment, such as Yokogawa, are mostly focused on the potential benefits of the product. Such benefits are offered to their customers, in a wide range of industries, thanks to the integration of IIoT technologies into the business. Although a large number of technologies necessary for adequate and flawless functioning as well as for IIoT support are already available, the necessary infrastructure is still not fully implemented in industrial facilities. To reduce the company's costs for engineers and designers, as well as the cost of their travel, and to compensate for lost time, Yokogawa is developing effective and secure methods to realize remote collaborations between engineers and designers via the Internet. Such ability to communicate and exchange data and ideas would allow engineers to work in teams, through web applications, in real time, on projects regardless of their current location.

Thus, for example, DSC (distributed control system) engineers located in the Yokogawa Industrial Automation (IA) center in Tokyo were able to effectively cooperate with EPC (Engineering, Procurement and Construction) engineers from the USA, engineers of suppliers of packaged and finished goods in Europe, as well as with engineers and end users from a Middle East ownership company. By accurately analyzing the data obtained, Yokogawa identified the great advantages of this type of distance cooperation:

- Not every product needs to be physically delivered
- In the future, there is the possibility of conducting testing on multiple components via the Internet
- Save time and costs for engineers and companies
- The re-testing process has been accelerated, as well as the process of final checking the functionality of equipment and software, and assuring them of their complete correctness before delivery.
- By analyzing the obtained data and storing them, information is available to end users 24 hours a week in order to resolve daily and extraordinary business operations
- Also, large amounts of data are generated from connected devices that serve as a form of database that will make it easier for each user to work as follows:
  - Over time, the amount of data obtained will grow
  - New challenges will be created in operational technologies (OT) and information technologies (IT).

Yokogawa believes that analyzing OT data generated through IoT device tracking, in terms of the most sought-after troubleshooting tools, can help its end users find an easy, simple and fast way to solve their problem. Thanks to the nature of IoT devices, the ease of availability of the required data, and the connection of IoT devices to company assets, leading people can easily access business statistics and thus industrial organizations can improve their business at any stage of business and thus improve TCO (Total Cost). of Ownership) and achieve operational excellence (OpX - Operations Excellence). Yokogawa

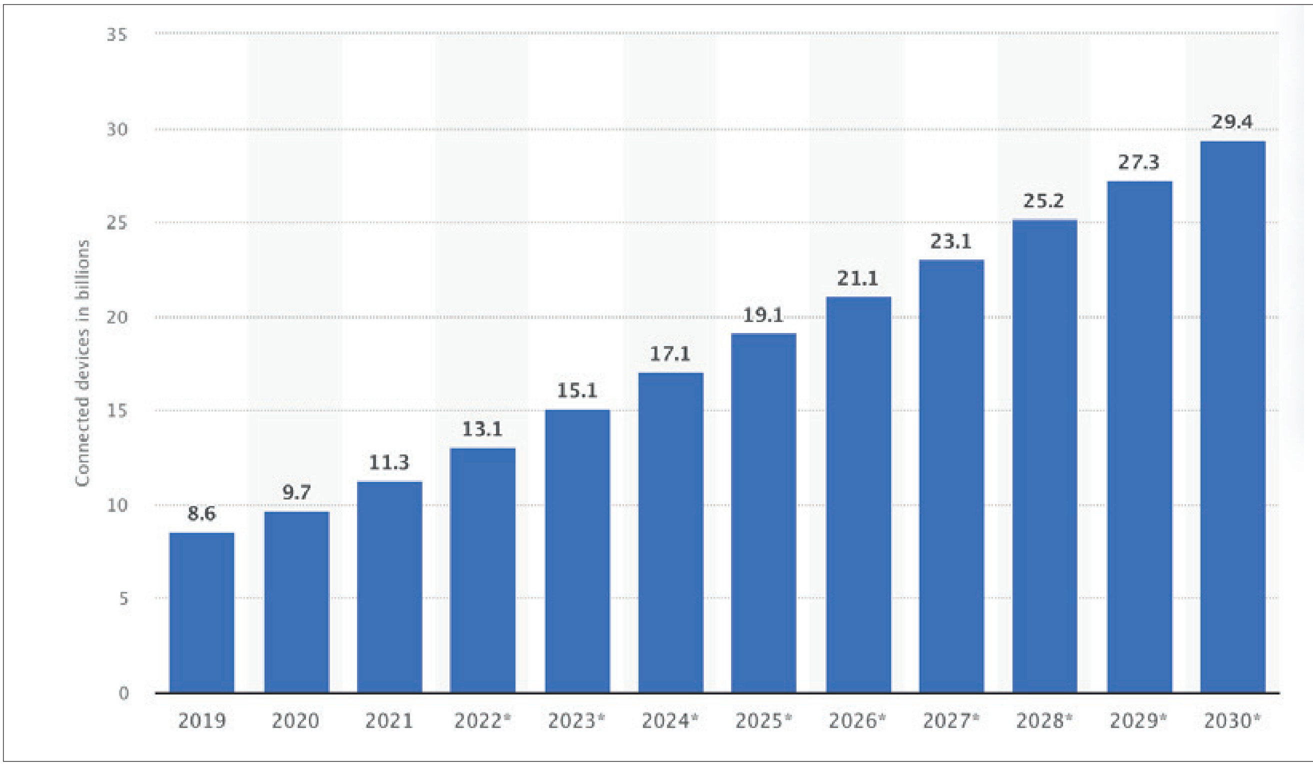


Figure 1: *Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2030 (in billions)/https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/ 4.4.2022.*

focuses its IoT business on two areas:

- Providing feedback to company professionals to interpret and act in extraordinary OT work processes
- Development of real-time logical and intelligent applications to trigger appropriate automated responses in structural process systems

The key success factors of IoT are:

- Appropriate knowledge in the field of OT and IT
- Clear understanding of operational requirements, which include the need for flexibility and extensibility to address
- Close integration of OT and IT
- Close cooperation between OT and IT suppliers, suppliers and end users, plant engineers as well as process engineers and data collection and processing experts all in order to facilitate and speed up the work process
- Carefully consider how the solution will be maintained and supplemented over time
- Ability to adapt to new business methods
- Impenetrable virtual security
- Secure and durable networks (wired and wireless).

Intel Approach

On the other hand, Intel believes that the opportunities that IoT creates also face great challenges, both for user IoT and for industrial IoT, whose operation and use are much more complex. The challenges that users face when integrating IoT technologies into their business are price, security, organizational alignments and changes in business processes.

Intel's IoT standards help integrate the core platform with the following benefits:

1. Integration from end to end, i.e. the existence of only one IoT network in which all IoT devices of one company are harmoniously integrated. The advantage of such a network system is the reduction of costs and business complexity.

2. Multiple solutions for one problem or user task, i.e. the same interest groups have the opportunity to place products on the market faster.

3. Improved security of end users and their data.

4. Better product quality as well as open standards expand the range of possible solutions from which the user chooses the best for their business.

5. Reducing financial costs and increasing the efficiency of the company.

The three key problems of integrating IoT technologies into business are:

1. Initial financial costs are difficult to justify investments (large initial investment, long process of return on invested finances and time)

2. IoT changes already existing business processes (it is necessary to educate a large number of employees, change the structures and devices necessary to maintain the work process, ie a long and financially demanding transition process)

3. IoT requires uniformity and harmonization of the way of working work of all companies that use IoT technologies.

Tasks and obligations of the company's IT sector in the integration of IoT technologies into business include:

- Business security and the establishment and maintenance of networks used by IoT devices for their work process
- Use of existing infrastructure
- Limited Internet access, which means targeted access, i.e. the worker can use only the tools necessary for the operation of IoT technologies while other tools are not available to the user
- Control of output data, i.e. what is available to the end user
- Care and maintenance of security and protection systems in order to maintain the confidentiality of company and user data
- Processes for suppliers:
    - Compactness of sensors as well as universality of devices and technologies, i.e. that IoT technologies can be used on all smart devices regardless of the year of production
    - Open standards (software that easily shares data with applications and thus allows the user quick and easy access to the problem and solution in a few clicks)
    - Adaptability to user needs as well as easy technology management
    - Security requirements in the form of obtaining a confirmation of receipt of the order as well as who is responsible for receiving the order or service.

After years of setting up individual IoT solutions and overcoming many obstacles and challenges that arose in meeting new workflows and tasks, and careful analytical work in many key segments, Intel has developed several standard processes for implementing IoT in its business:

- Defined IoT system based on user needs and processes to be monitored
- Classification of data obtained from smart device sensors, ie creating a secure infrastructure that will enable proper management, analysis, protection and storage of data necessary to solve given problems
- Designing a secure and accurate network infrastructure
- Compliance with the rules on corporate data management, ie care for the confidentiality of company data
- Integration and management of IoT devices
- Designing IoT systems so that the work process runs smoothly and easily, which is actually the goal of introducing IoT
- Establishing an end-user support model to quickly diagnose and resolve issues.

Company Approach

The company approach to IoT technologies and their implementation in business processes includes several interrelated processes:

- Development of company-specific IoT, i.e. for the needs of the company's work process in order to accelerate the solution of possible problems
- IoT protection plan and thus confidential data necessary for operation
- IoT device and end user communication plan
- Value-added analytics and the creation of useful algorithms that would speed up the work process and easily overcome possible problems in the work process itself, but there is a chance of too much intrusion into the company's private data
- The diversity of IoT applications and their enormous demand for simplicity represent ideal business and financial opportunities and challenges for IT companies and professionals
- The economic potential or financial savings of the company from the IoT are estimated at 11.1 billion dollars per year until 2025.
- The diversity of applications and devices poses a challenge to the IT world (how to imple-

ment IoT technologies in the business environments of most if not all companies).

Through the implementation of IoT technologies in work processes, large global companies have discovered a number of benefits, the most significant of which are:

- Identifying the needs of companies based on the use of standardized algorithms and accurate imitation of the workflow on the IoT device
- Detailed development of a database of possible problems and the simplest solutions to these problems and tasks for the same stakeholders
- Mapping or merging the most important processes into a functional component that can be presented to basic microservices and IoT resources and thus provide a link to OT
- Organizing and executing harmonized business processes taking into account the flow of large amounts of data as well as the processes of IoT devices themselves that are independent of users, and which are necessary for the precise functioning of the system.

Currently, there is no fully developed IoT system with all the components, but there are a number of bases for such a system (which can be used to obtain a system that is at the same time easy to use, fast, reliable and secure). Work should be done to reduce the gap between the company's perspectives and IoT-focused processes and thus enable applications to be more accessible and uniform for companies and end users in order to realize the full potential of IoT technologies and improve the quality of workflow with financial savings.

Planning and development of business processes involves identifying and defining patterns of business processes within the company or in parts of the company, which can be modernized, and these business patterns include different sources of data and tools for doing business processes. The focus is on processes involving physical means and machines and the ways in which they are related to IoT technologies. In this way, it is possible to analyze and optimize business and introduce fully automated business processes in the work of the company. The highest intensity in automation will be held by the following: production, smart buildings, utility parts of the company, agricultural activities and process industry.

Business Process Model and Notation (BPMN) is a specific standard developed to improve and facilitate business processes. Thus, formal specific processes can be produced that can be executed and then mapped to microservices and IoT resources. BPMN provides companies with the ability to understand and communicate internal business processes in graphic markup based on XML (Extensible Markup Language). Furthermore, in this way, organizations are enabled to quickly adapt to new internal and B2B (Business to business) business needs and circumstances.

IoT operational models are equipped with a basic platform to enable adequate operation of applications and their tools. To date, these platforms have mainly focused on infrastructure services: connectivity, protocol customization, device management, and secure device shutdown. Aspects of the platform are still largely focused on data, i.e. their collection and storage, while the use of this data in the business process is still in its infancy and the full potential of IoT technologies, as originally conceived, has not yet been realized. IoT applications can be enriched and expanded with significant capabilities, as well as effectively integrated into the company's complex business processes. Integration of IoT applications in business processes takes place in several significant phases that require knowledge, time and finances. The basic things to keep in mind are as follows:

- IoT applications should not be viewed in isolation or as an isolated solution to specific and limited problems, but as a set of useful data and procedures that will lead us to the desired goal in the easiest way
- IoT devices should be viewed as parts of the company and it is necessary to integrate the obtained data into the overall processes of the company
- The approach to business process modeling should ensure that the implementation of IoT technologies is efficient and to enable full automation of the company.

In addition to the above company approaches of Yokogawa and Intel, Cloud is also very important for the progress of IoT. Considering the strength of the connection between these two technologies, the concept of CloudIoT was created, which represents the integration of Cloud and IoT. With Cloud computing, resources, applications, and information shared over the Cloud are also available to other computers and smart devices connected to the Internet. Also, the term Big Data is very interesting, referring to technology that collects and analyzes large amounts of data in real time, but they have a big disadvantage that they collect different data, different accuracy, from both reliable and unreliable sources. Big Data and IoT are closely connected, because with the help of devices and sensors, they collect extremely large amounts of data that are very useful, easily accessible to everyone and have some significance for the organization.

Today we have quite a few things that represent some kind of IoT technology. The Tesla car is at the forefront of this. Namely, this is an electric car that is completely connected to the Internet, collects, analyzes and transmits data in real time, just as the IoT device should work. Also, e-health, which is slowly being introduced in every health system, is a good example of the application of IoT technologies, because health workers have real-time access to a large amount of data that allows easier, faster and more efficient business process of a healthcare institution.

In addition to the company's approach to IoT, it is important to mention the commercial approach to IoT technologies. The goal is to create a smart home that, with the help of smart devices connected to the Internet, makes a home tailored to the owner, ie according to all his life habits. In a smart home, the light is turned on via an application through which the quality of lighting, colors as well as the moment when the light is turned on and off can be controlled. In addition to light bulbs, smart security camera systems, sockets, switches, cooling/heating bodies, etc. are in use today, which are connected via a large number of sensors to an application through which their operation can be controlled.

Unlike corporate IoT, commercial IoT is much simpler, devices are cheaper and more affordable, and easily replace existing devices. One of the most common commercial IoT devices is SMART TV, which is connected to the Internet and is a kind of IoT device. In order for it to become an IoT device in a narrower sense, thanks to technological advances, it is now possible to give commands to the TV by voice, and the latest way of management is that the TV recognizes user gestures and therefore takes certain action. When it comes to controls, the TV monitors and records TV content that the user likes to watch and suggests similar content. All this data is collected, processed and a database of users is created, and this data is further sent to TV manufacturers as well as to companies that offer TV content broadcasting.

In 2021, the first smart bus station was set up in Banja Luka, which also represents the commercial use of IoT technologies. The advantage of such a station is the availability of the current timetable (this has not been the case so far), it provides insight into the map of bus stops as well as data on the current location of the desired bus. The smart station offers the possibility of connecting to the Internet, and thus performs data collection, storage and analysis of data.

Conclusion

Taking into account all the data known so far, the use of IoT technologies to improve the work process has many advantages, but also disadvantages, which are probably crucial why IoT technologies have not yet been implemented in business. Out of the many advantages, first of all, the most significant is the acceleration of the work process through automation with minimal error in business. IoT allows us to solve many requests very accurately and efficiently in a short time. Furthermore, since smart devices replace the human workforce, it is possible to perform the work process continuously, 365 days a year. Another in a series of advantages is that IoT, in fact, represents a kind of base of potential solutions to any problem that may be encountered in business. The solutions are classified by categories, are universal, and according to IoT principles are available to their users 24 hours a day. The last, but not less important advantage, is the ability to analyze the work process at any time, identify potential errors and eliminate them in the shortest possible time, all in order to improve business.

The disadvantages of IoT technologies and their implementation in business processes are, primarily, that there is a gradual reduction in labor demand (because at the heart of technology is a smart machine that analyzes and sends the necessary data).

Also, a lot of confidential information about the company's business, private data about workers, machines and production processes becomes available to other users of IoT technologies and as such can be used illegally. The personal stamp of the company, which is reflected in the specifics of the company's work process, is eliminated, because most of the work within IoT technologies is carried out according to given algorithms and there is no room for creativity and creating some new solutions. In theory, the IoT is expected to operate the device flawlessly for 365 days, 24 hours a day, without user correction or service intervention. However, imperfections in the design, testing, and all other features involved most often lead to several failures during the operation of the IoT device.

REFERENCES

[1] W. E, Building the internet of things using RFID: the RFID ecosystem experience, IEEE Internet Computing, 2009.

[2] N. A, Internet of things enabled, 2016.

[3] [Online]. Available at: 1. https://www.ericsson.com/en/blog/2020/6/enterprise-perspective-iot-modeling . [Accessed 12 5 2022].

[4] [Online]. Available at: https://www.switchautomation.com/enterprise-iot/. [Accessed 13 5 2022].

[5] [Online]. Available at: https://internetofthingsagenda.techtarget.com/essentialguide/Framing-your-enterprise-IoT-approach . [Accessed 14 5 2022].

[6] [Online]. Available at: https://internetofthingsagenda.techtarget.com/essentialguide/Framing-your-enterprise-IoT-approach . [Accessed 15 5 2022].

[7] [Online]. Available at: https://techinformed.com/internet-of-things-key-stats-for-2022/. [Accessed 3 5 2022].

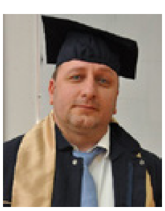
[8] A. Z. Čabarkapa M., "Road safety management in local communities," *JITA – Journal of Information Technology and Applications, Banja Luka, Pan-Europien University APEIRON, Banja Luka, Republika Srpska, Bosna i Hercegovina, ,* vol. 9, no. 2:12-34, 2019.

List of abbreviations and acronyms

BPMN – Busniess Process Model and Notation
B2B – Business to Business
CloudIoT – Integration of Cloud and IoT
DSC – Distributed Control System
EPC – Engineering, Procurement and Construction
IA – Industrial Automation
IIoT – Industial Intertnet of Things
IoT – Intertnet of Things
IT – Information Technologies
OpX – Operations Excellence
OT – Operating Technologies
TCO – Total Cost of Ownership
XML – Extensible Markup Language

Received: May 18, 2022
Accepted: May 28, 2022

ABOUT THE AUTHORS



Dražen Marinković was born on 1978 in Banjaluka. He finished elementary school and high school in electrical engineering. The Faculty of Business Informatics at the Pan-European University APEIRON Banjaluka enrolled in the academic year 2009/2010. years and graduated with an average grade of 9.81, defending a master's thesis entitled: "NoSQL databases in theory and practice" with a grade of 10. Academic 2015/2016 enrolls in doctoral academic studies of the third cycle at the Pan-European University APEIRON in Banjaluka, at the Faculty of Information Technology, and in the doctoral program "Information Systems in Communications and Logistics" where he passed all exams with excellent grades - excellent, 10. During his doctoral academic studies, as an author and co-author, he published 15 scientific and professional papers, two of which were published in international peer-reviewed journals.

Zoran Avramović was elected a full member of the Russian Academy of Transport (RAT, St. Petersburg, Russia, since 1995), the Russian Academy of Natural Sciences (RAEN, Moscow, Russia, since 2001), the Serbian Academy of Engineering (IAS, Belgrade, since 2004). (formerly: Yugoslav Academy of Engineering - JINA) and the Academy of Electrical Engineering of the Russian Federation (RAEN, Moscow, Russia, since 2007). He is the scientific secretary of the Electrical Engineering Department of the Engineering Academy of Serbia. So far Avramović has published: 371 scientific and professional papers.

FOR CITATION

Dražen Marinković, Zoran Ž. Avramović, IoT – Company Approach to IoT Modeling and Applications, *JITA – Journal of Information Technology and Applications, Banja Luka*, Pan-Europien University APEIRON, Banja Luka, Republika Srpska, Bosna i Hercegovina, JITA 12(2022) 1:48-54, (UDC: 621.391:004.738.5), (DOI: 10.7251/JIT2201048M), Volume 12, Number 1, Banja Luka, June (1-64), ISSN 2232-9625 (print), ISSN 2233-0194 (online), UDC 004

Encapsulation and Functionality of Sensor Systems in the Welding Process

Barbara Bagarić

Panevropski univerzitet APEIRON Banja Luka, Republika Srpska, BiH, barbara.bagaric2@apeiron-edu.eu

Case study

https://doi.org/10.7251/JIT2201055B UDC: 007.52:[681.586.74:517.938

Abstract: This paper shows some aspects of interaction between human and robots and role of the sensors in welding process. Today, use of the robots is very important in modern industry. Sensors are very important in welding process and they increase the productivity and precision of the robot. It is very important to optimize use of the sensor, from choosing the right programming metod, creating good environment such as acceptable amounts of humidity in the air, temperature to educating the operators to be able to work with robot machine and understand and run pre-programmed codes. When all the mentioned steps are correctly defined, sensor will work perfectly and production can be as good as possible.

Keywords: robots, online programming, offline programming, welding process.

Introduction

Today, robots have taken a very important position in modern industry. Human-robot interaction plays a significant role in understanding, developing and evaluating the systems that humans control or collaborate with. Before starting to use robots, it is important to consider the basic principles and methods of robot operation, as well as the history of human interaction with robots. Also very important are the relationships between humans and robots when collaborating, and the challenges we face when interacting. In this paper, the use of robots and sensor systems in the welding process is processed and studied, with an emphasis on the use of iCSE sensor system and the functionality of offline robots. It is also very important to mention the importance of robotics, both in industry and welding technology, the basics of robot systems used in industry, welding robot programming, sensors to control the welding process, weld monitoring, robotic process control and the role of sensor integration monitoring the degree of adaptability on the example of an offline robotic welding system.

The practical part of the work, which was performed in the company ThermoFLUX with its headquarters in Jajce, in the department of the welding plant on the CLOOS robot, samples were made in the welding process.

Robots are machines that are capable of automatically performing a series of complex operations and actions. They do it very precisely, quickly and almost without error. Robots are programmed using computers and programming languages. Most often, robots are guided by an external control device, which is also called a joystick, but they can also be controlled by controls that are built into the robot. Robots can be constructed in a way that evokes human form, while most robots are machines designed to perform tasks, designed with an emphasis on functionality, not so much on aesthetics and imitation of human appearance. Robots can be:

- Autonomous
- Semi-autonomous
- Humanoids

Basics Of Industrial Robots

Industrial robots are a combination of a robotic system whose main role is to participate in the production process. Such robots are automated, programmed and capable of moving along several axes. Industrial robots have found the greatest application in the processes of welding, assembly, disassembly, varnishing, selection and installation of various parts, packaging and labeling of products and the like, all for the purpose of more durable, faster and more accurate manufacturing process. It is currently estimated that there are about 1.60 million industrial robots in the world in daily operation worldwide.

Welding Robot Programming

Online Programming Method

The online programming method is also called the network programming method and includes stopping the robot, putting it out of production and putting it into programming or learning mode. Programs are then created or modified using the joystick.

Joystick programming is the most common method for programming industrial robots. It is estimated that about 90% of industrial robots are programmed through the same. The joystick is also called the robot motion control control box. It's also called a "learning box". The robot is set up in a "learning" or "teaching" manner, and the pendant is used to control the robot step by step. The robots come with a compatible device that connects to the robot. Developers use the pendant interface to create and modify programs. Online robotic programming is a method in which a robotic arm is guided exactly through the waypoints of an application.

Offline Programming Method

Offline programming is computer-based programming that does not require robot movement, so it is convenient because the production process does not have to stop. The purpose of such programming is to create as many technological processes as possible on one robot and its workplace.

Offline programming includes the creation of programs via a computer using programming software and a simulation work environment. Unlike online programming methods, offline robotic programming does not require the presence of a robot, nor its exclusion from the production process if the robot is already set up in production. The robot can stay in operation while users create or modify applications using offline software. The simulation creates 3D replications of the robot's work environment so that thorough testing and debugging of the resulting programs can be performed. When the program is ready, it is downloaded to the robot and continues with production.

iCSE sensor system is a non-contact measuring system that works on the basis of laser triangulation. In this process, changes in the distance between the sensor and the reflecting surface are determined.

- Relative position measurements can be performed with the system.
- Furthermore, geometry readings from the sensor can be used as reference measurements.
- The laser beam of the system must not be aimed at people.

The following factors may affect proper operation:

- Placing untrained personnel on the sensor system. Failure to follow the instructions for use.
- Any other use not considered intended.
- Failure to follow work instructions on site.
- External factors
- Dirt on the lens
- Lens damage
- Oil film on the lens
- Irradiated with UV light
- Vapors from welding in the sensor beam path
- Direct sun
- Reflective surfaces on the measuring object
- Reflections from mechanically treated surfaces on the measuring object
- Dirty protective cap

Technical Properties of Icse Sensor System

According to the book: PA_iCSE_Rev. 3.1 :

Sensor is a device that measures certain parameters in the environment of the robot and allows the robot to adapt to the conditions in which it is at a certain moment. Accordingly, in robotic welding,

TECHNICAL PROPERTIES OF ICSE SENSOR SYSTEM									
Dimensions	Housing	Weight	Operating temperature	Humidity	Nominal measuring distance	Measuring range	Vertical resolution	Horizontal resolution	Beam diameter
H = 160 mm W = 85 mm D = 42 mm	Aluminum, anodised (black)	approx. 630 g	0 - 50 ° C	35 - 80%	100 mm	+/- 40 mm	16 µm	50 µm	0.1 mm

TECHNICAL PROPERTIES OF ICSE SENSOR SYSTEM									
Reaction time	Power adjustment	Output signal	Tracking speed	Min. edge height	Min. spacing width	Search accuracy	Laser diode power	Laser diode wavelength	Laser protection class
2.5 ms	a car	+/- 4V	approx. 300 cm / min	1 mm	1 mm	+/- 0.2 mm (depending on surface, speed, material, etc.)	max. 1 mW	650 nm, visible	2 according to DIN EN 60825-1"[1]

sensors are used to measure parameters in order to achieve a process that meets the required requirements in the best and most economical way.

Application of Sensors In Robotic Welding

A sensor is a device that measures certain parameters in the environment of the robot and allows the robot to adapt apropos to the conditions in which it is at a certain moment. Accordingly, in robotic welding, sensors are used to measure parameters in order to achieve a process that meets the required standards in the best and most economical way.

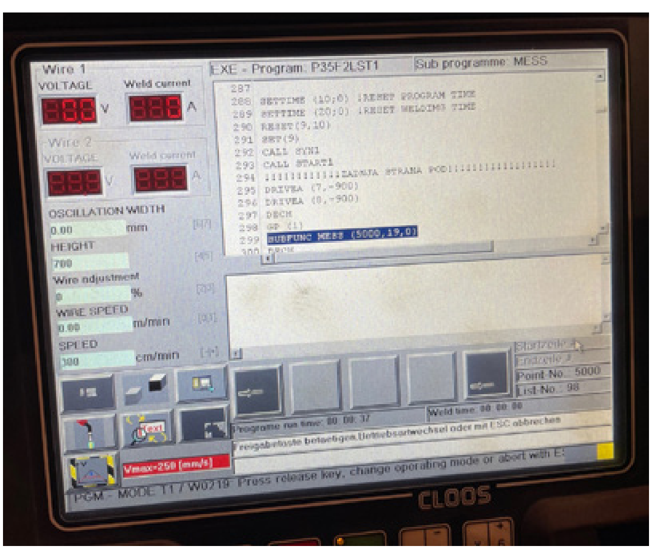


Figure 1 *Template 19*

Before robot starts the welding process and understanding the sensor function, we will explain the templates according to which the welding is performed. According to Template 19, which is used for welding the back or floor of the boiler. The program was entered into the robot by the offline method of programming and saved as Template 19. Speed of robot metric system in this template is 300 cm in 1 minute. After the welding process is over robot will take oscillations in width and height. At beginning of the process, oscillation in width and height is 0.00 mm, Wire adjustment is 0% and Wire speed is 0.00 mm.

Program:

!!!!!!!!!!!!!!ZADNJA STRANA!!!!!!!!!!!!!!

DRIVEA (7, -900)

DRIVEA (8, -900)

DECH

GP(1)

SUBFUNC MESS (500, 19, 0) - This part of the code in the program is in charge of calling Template 19 to be activated. After this part of the code, the robot starts the welding process according to the entered parameters by the template.

After the welding process with the existing Template 19, we transfered the robot to the measurement and later welding process. Figure 1 shows the path taken by the robot when making the weld. When programming the robot on template 19, the path taken by the robot is plotted.

[1] PA_iCSE_Rev. 3.1, Carl Cloos Schweisstechnik GmbH

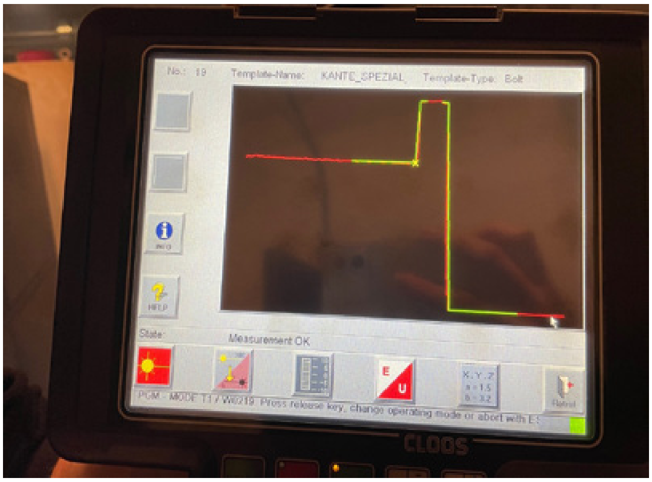


Figure 1 *Template 19 - new measurement*

The red line shows the plotted path while the green line shows the actual path that the robot went through while welding. We can notice that the difference in the path of the weld entered during programming and the actual, performed path is almost imperceptible, and we conclude that the sensor in this case correctly measured the path by the robot arm. The unit of deviation during the measurement was within the allowed limits, so the robot continued with the welding process. Given the large number of external factors that in this case could have caused damage during the operation of the sensor, which, among other things, partially affected the measurement of the sensor, it can be concluded that in this case the sensor performed the correct measurement. After comparing the programmed trajectory with the actual one and determining a small deviation within the limits, the robot starts its welding on the boiler.

Figure 2 *Deviations from the entered and measured path*

Figure 2 shows deviations from the entered and measured path. The template and its path are activated, the laser head has performed the path measurement process. The position according to the Y coordinate is 30.45 mm, according to the Z coordinate 82.40 mm.

Figure 3 shows the oscillations in the sensor measurement. According to figure 1 we can see the difference in the oscillation width and height, Wire adjustment and Wire speed. These are all values that the robot has recorded during the measurement and welding process. Oscillation in width is 1.58 mm, Wire adjustment is 2% and Wire speed is 13.50 m/min. The results are shown in the table 1:

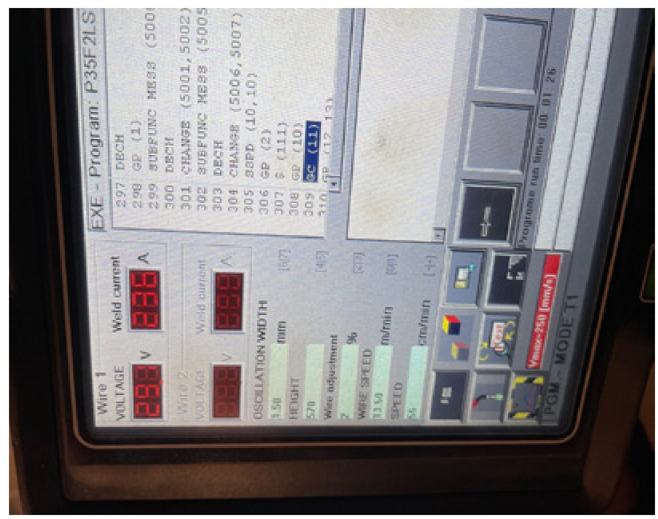


Figure 3 *Oscillations*

CONCLUSION

Encapsulation and functionality of sensor systems in the welding process is important class in the area of robot welding. Their essential function is to provide a safe space near the robot, and also to ensure fast, efficient and functional work. Before starting to use a robot, it is important to consider its principles and methods of work, as well as the way in which the robot interacts with the human environment. This paper investigates and presents how the sensor works, how it measures and determines the path by which it performs the welding process and what external harmful consequences can affect the quality of its work. From the research, we conclude that the robot made only small oscillations during welding according to the template entered

Table 1 *Oscillations at measured and default values*

	Position Y	Position Z	OSCILLATIONS (mm)	Wire adjustment (%)	Wire speed (m/min)	Speed (cm/min)	VOLTAGE
Default values	28,95	80,9	0	0	0	300	0
Measured dimensions	30,45	82,4	1,5	2	13,5	55	29,1

by the offline programming method, which did not significantly affect the welding process and its quality. The oscillations are within the allowable limits. It can be seen from the graph that the measured oscillation is 1.50 mm.

It is precisely these processes that require rapid change, which makes it very difficult to find literature that corresponds to this type of research. The information provided and explained in the paper are from the factory where heat boilers and pellet fireplaces are produced and is the result of practical work. It is precisely these processes that require rapid change, which makes it very difficult to find literature that corresponds to this type of research. The information provided and explained in the paper itself is from the factory where heat furnaces are produced and is the result of practical work.

References

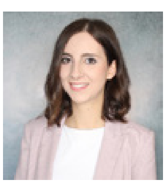
[1] PA_iCSE_Rev. 3.1, Carl Cloos Schweisstechnik GmbH, 10/2017 Rev. 3.1

[2] PA_QIROX_V08.xx_Rev. 01/17_EN, Carl Cloos Schweisstechnik GmbH, 22.03.2017 Rev 01/17

[3] https://www.cloos.de/de-de/, 2022 Carl Cloos Schweisstechnik GmbH

Received: January 12, 2022.
Accepted: May 27, 2022.

About the authors



Barbara Bagarić (1999) was born in Jajce, Bosnia and Herzegovina. She finished elementary school in Jajce. She finished secondary school of Gimnasium in Jajce. She defended her bachelor's degree at the Faculty of Information Technology, Pan-European University "Apeiron" in Banja Luka, majoring in Computer Multimedia and Graphic Design, and Information Techology. She is currently studying for a master degree under the mentorship of prof. dr. Željko Stanković. She is a teacher of informatics at the primary school, a teaching assistant at the Pan-European University "Apeiron", Banja Luka and the founder of the Photo & Design Studio Seven based in Jajce.

For citation

Barbara Bagarić, Encapsulation and Functionality of Sensor Systems in the Welding Process, *JITA – Journal of Information Technology and Applications, Banja Luka*, Pan-Europien University APEIRON, Banja Luka, Republika Srpska, Bosna i Hercegovina, JITA 12(2022) 1:55-59, (UDC: 007.52:[681.586.74:517.938), (DOI: 10.7251/JIT2201055B), Volume 12, Number 1, Banja Luka, June (1-64), ISSN 2232-9625 (print), ISSN 2233-0194 (online), UDC 004

Instructions for Authors

The *Journal of Information Technology and Application (JITA)* publishes quality, original papers that contribute to the methodology of IT research as well as good examples of practical applications.

Authors are advised that adherence to the Instructions to Authors will help speed up the refereeing and production stages for most papers.

- Language and presentation
- Length of submissions
- Submission
- Contact details/biographies
- Title of the paper
- Abstract and keywords
- Figures and tables
- Sections
- Footnotes
- Special characters
- Spelling
- References
- Proofs
- PDF offprint
- Copyright and permissions
- Final material
- Correspondence
- Publication ethics

Language and presentation

Manuscripts should be written in English. All authors should obtain assistance in the editing of their papers for correct spelling and use of English grammar. Manuscripts should have double spacing, with ample margins and pages should be numbered consecutively. The Editors reserve the right to make changes that may clarify or condense papers where this is considered desirable.

Length of submissions

Papers should not normally exceed 12 Journal pages (about 8000 words). However, in certain circumstances (e.g., review papers) longer papers will be published.

Submission

Manuscripts must be submitted through the JITA online submission system.

Please read the instructions carefully before submitting your manuscript and ensure the main article files do not contain any author identifiable information.

Although PDF is acceptable for initial submission original source (i.e. MS Word) files will be required for typesetting etc.

Contact details/biographies

A separate file containing the names and addresses of the authors, and the name and full contact details (full postal address, telephone and e-mail) of the author to whom correspondence is to be directed should be uploaded at the time of submission (you should select Contact details/Biographies as the file type). This file is not shown to reviewers. This file should also contain short biographies for each author (75 words maximum each) which will appear at the end of their paper.

The authors' names and addresses must not appear in the body of the manuscript, to preserve anonymity. Manuscripts containing author details of any kind will be returned for correction.

Title of the paper

The title of the paper should not be longer than 16 words.

Abstract and keywords

The first page of the manuscript should contain a summary of not more than 200 words. This should be self-contained and understandable by the general reader outside the context of the full paper. You should also add 3 to 6 keywords.

Figures and tables

Figures which contain only textual rather than diagrammatic information should be designated Tables. Figures and tables should be numbered consecutively as they appear in the text. All figures and tables should have a caption.

Sections

Sections and subsections should be clearly differentiated but should not be numbered.

Footnotes

Papers must be written without the use of footnotes.

Special characters

Mathematical expressions and Greek or other symbols should be written clearly with ample spacing. Any unusual characters should be indicated on a separate sheet.

Spelling

Spelling must be consistent with the Concise Oxford Dictionary.

References

References in the text are indicated by the number in square brackets. If a referenced paper has three or more authors the reference should always appear as the first author followed by et al. References are listed alphabetically. All document types, both printed and electronic, are in the same list. References to the same author are listed chronologically, with the oldest on top. Journal titles should not be abbreviated.

- Journal

[1] Avramović Z.Ž. (1995). Method for evaluating the strength of retarding steps on a marshalling yard hump. *European Journal of Operational Research*, 85(1), 504–514.

- Book

[2] Walsham G. (1993). *Interpreting Information Systems in Organizations*. Wiley, Chichester.

- Contributed volume

[3] Huberman A.M. and Miles M.B. (1994). Data Management and analysis methods. In *Handbook of Qualitative Research* (Denzin NK and Linkoln YS, Eds), pp 428-444, Sage, Thousand Oaks, California.

- Conference Paper

[4] Баранов Л.А. (2017). Принципы построения и алгоритмы интеллектуальных автоматических систем управления движения поездов, функционируемых в рамках городских транспортных систем. In G. Radić & Z.Ž. Avramović (Eds.), *Proceedings of a IX International Scientific Conference "Information Technology for e–Education"*, (pp.9-18). Pan-European University APEIRON, Banjaluka, 29–30.9.2017. Republic of Serpska, B&H

- Unpublished reports/theses

[5] Nandhakumar J.J. (1993). The practice of executive information systems development: and in-depth case study. *PhD Thesis*, Department of Engineering, University of Cambridge.

Proofs

Proofs of papers will be sent to authors for checking. Alterations to diagrams should be avoided where possible. It will not be possible to accept major textual changes at this stage. Proofs must be returned to the publishers within 48 hours of receipt by fax, first-class post, airmail or courier. Failure to return the proof will result in the paper being delayed.

PDF offprint

Corresponding authors will receive a PDF of their article. This PDF offprint is provided for personal use. It is the responsibility of the corresponding author to pass the PDF offprint onto co-authors (if relevant) and ensure that they are aware of the conditions pertaining to its use.

The PDF must not be placed on a publicly-available website for general viewing, or otherwise distributed without seeking our permission, as this would contravene our copyright policy and potentially damage the journal's circulation. Please visit http://www.jita-au.com to see our latest copyright policy.

Copyright and permissions

The copyright of all material published in the Journal is held by Pan-European University APEIRON. The author must complete and return the copyright form enclosed with the proofs.

Authors may submit papers which have been published elsewhere in a foreign language, provided permission has been obtained from the original publisher before submission.

Authors wishing to use material previously published in JITA should consult the publisher.

Final material

All final material must be submitted electronically in its original application format (MS Word is preferred). The file must correspond exactly to the final version of the manuscript.

Correspondence

Business correspondence and enquiries relating to advertising, subscriptions, back numbers or reprints should be addressed to the relevant person at:

Pan-European University APEIRON
Journal JITA
Pere Krece 13, P.O.Box 51
78102 Banja Luka
Bisnia and Hercegovina / RS
E-mail: jita@apeiron-edu.eu

Publication ethics

We take an active interest in issues and developments relating to publication ethics, such as plagiarism, falsification of data, fabrication of results and other areas of ethical misconduct. Please note that submitted manuscripts may be subject to checks using the corresponding service, in order to detect instances of overlapping and similar text.

JITA

Journal of Information Technology and Applications
Pan-European University APEIRON Banja Luka
www.jita-au.com
jita@apeiron-edu.eu

PUBLISHER: **Pan-European University APEIRON**, Banja Luka
College of Information Technology Banja Luka, Republic of Srpska, B&H
www.apeiron-uni.eu
Darko Uremović, Person Responsible for the Publisher
Aleksandra Vidović, PhD, Editor of University Publications

ISSN 2232-9625
9 772232 962005