

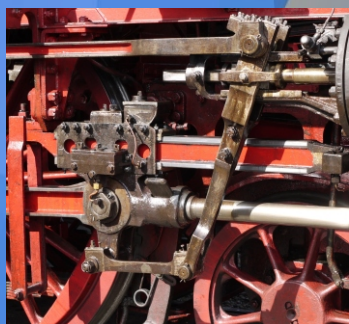
INTERNATIONAL SCIENTIFIC JOURNAL

INDUSTRY 4.0

ISSN (PRINT) 2534-8582

ISSN (WEB) 2534-997X

YEAR V ISSUE 2/2020



①



②



③



④

PUBLISHED BY
SCIENTIFIC TECHNICAL UNION OF MECHANICAL ENGINEERING
“INDUSTRY 4.0”, BULGARIA

INTERNATIONAL SCIENTIFIC JOURNAL

INDUSTRY 4.0

YEAR V, ISSUE 2 / 2020
ISSN (PRINT) 2543-8582, ISSN (WEB) 2534-997X

PUBLISHER
**SCIENTIFIC TECHNICAL UNION OF MECHANICAL
ENGINEERING “INDUSTRY 4.0”**

108, Rakovski Str., 1000 Sofia, Bulgaria
tel. (+359 2) 987 72 90,
tel./fax (+359 2) 986 22 40,
office@stumejournals.com
WWW.STUMEJOURNALS.COM

EDITOR IN CHIEF

Prof. D.Sc. Georgi Popov, DHC, Technical University of Sofia, BG
Prof. Dr. Dr. Jivka Ovtcharova, DHC, Karlsruhe University of Technology, DE

EDITORIAL BOARD

Members:

Acad. Igor Bychkov, RU
Cor. member Alexey Beliy, BY
Cor. member Svetozar Margenov, BG
Prof. Alexander Afanasyev, RU
Prof. Alexander Guts, RU
Prof. Andrzej Golabczak, PL
Prof. Andrey Firsov, RU
Prof. Bobek Shuklev, MK
Prof. Boris Gordon, EE
Prof. Branko Sirok, SI
Prof. Claudio Melchiorri, IT
Prof. Cveta Martinovska, MK
Prof. Dale Dzemydiene, LT
Prof. Dimitar Yonchev, BG
Prof. Dimitrios Vlachos, GR
Prof. Dragan Perakovic, HR
Prof. Galina Nikolcheva, BG
Prof. Galina Zhavoronkova, UA
Prof. Gerard Lyons, IE
Dr. Giovanni Pappaletta, IT
Prof. Henrik Carlsen, DK
Prof. Idilia Bachkova, BG
Prof. Idit Avrahami, IL
Prof. Inocentiu Maniu, RO
Prof. Iurii Bazhal, UA
Prof. Jürgen Köbler, DE
Prof. Jiri Maryska, CZ
Prof. Katia Vutova, BG
Prof. Lappalainen Kauko, FI
Dr. Liviu Jalba, RO
Prof. Luigi del Re, AT
Prof. Majid Zamani, DE

Prof. Martin Eigner, DE
Dipl.-Kfm. Michael Grethler, DE
Prof. Michael Valasek, CZ
Prof. Milija Suknovic, RS
Prof. Miodrag Dashic, RS
Prof. Mladen Velez, BG
Prof. Murat Alanyali, TR
Prof. Nafisa Yusupova, RU
Prof. Nina Bijedic, BA
Prof. Ninoslav Marina, MK
Prof. Olga Zaborovskaia, RU
Prof. Pavel Kovach, RS
Prof. Petar Kolev, BG
Prof. Peter Korondi, HU
Prof. Peter Sincak, SK
Prof. Petra Bittrich, DE
Prof. Predrag Dasic, RS
Prof. Radu Dogaru, RO
Prof. Raicho Ilarionov, BG
Prof. Raul Turmanidze, GE
Prof. René Beigang, DE
Prof. Rozeta Miho, AL
Prof. Sasho Guergov, BG
Prof. Seniye Ümit Oktay Firat, TR
Prof. Sreten Savicevic, ME
Prof. Stefan Stefanov, BG
Prof. Svetan Ratchev, UK
Prof. Sveto Svetkovski, MK
Prof. Tomislav Šarić, HR
Prof. Vasile Cartofeanu, MD
Prof. Vidosav Majstorovic, RS
Prof. Vjaceslavs Bobrovs, LV
Prof. Vladyslav Alieksieiev, UA

C O N T E N T S

TECHNOLOGICAL BASIS OF “INDUSTRY 4.0”

Data security model in cloud computing Ivan Kashukeev, Stoyan Denchev, Ivan Garvanov	55
CAD/CAM-system module for the design of automatic production Oleg Nikolaevich Mikhalev, Alexander Sergeevich Yanyushkin	59
Features of solving identification problems in transport technologies Anatolii Tymchenko, Volodymyr Boiko	63
A survey on deep learning in big data analytics Makrufa Hajirahimova, Aybeniz Aliyeva	68

DOMINANT TECHNOLOGIES IN “INDUSTRY 4.0”

Concept of parametric design and fast 3D printing of individual shoes Valentina Vladimirovna Kostyleva, Olga Borisovna Konovalova	72
An experimental study on energy generation from photovoltaic-thermal hybrid systems Danijela Kardaš, Bojan Knežević, Milan Pupčević	76

BUSINESS & “INDUSTRY 4.0”

Information risk management in SME sector enterprises Paweł Kobis	79
Introduction of Industry 4.0 in Industrial Enterprises: Problems and Challenges Neli Veleva	84
On predictability of precious metals towards robust trading Turan Erman Erkan, Adil Gürsel Karaçor	87

SOCIETY & „INDUSTRY 4.0”

Implementation of UAV for environment monitoring of a Smart City with an airspace regulation by AIXM-format data streaming Vladyslav Aliexsieiev, Bohdan Markovych	90
E-university cloud information security issues Firudin Aghaev, Gyulara Mamedova, Esmira Aleskerova	94
Personal data security techniques in e-education Firudin Aghaev, Gyulara Mamedova, Rena Melikova	97

Data security model in cloud computing

Ivan Kashukeev, Stoyan Denchev, Ivan Garvanov
University of Library Studies and Information Technologies, Bulgaria
ikashukeev@gmail.com, s.denchev@unibit.bg, i.garvanov@unibit.bg

Abstract: *In the last decade, cloud computing has become an extremely important business assistant that offers significantly lower costs than traditional computer-aided resource provision and greater adaptability to business changes. Thus, it helps them to achieve their business goals. The development of cloud computing poses significant risks, which serve as a partial barrier to their use by the businesses. Security and protection of information are considered to be one of the most critical points in the use of cloud computing. Compared to traditional solutions, cloud computing moves application software and databases to large data centers where information and service management are not always secure and reliable. As a result of this, the community and businesses have many concerns and questions about the security of data and applications in cloud infrastructures. In addition, concerns about data security and applications also arise due to the fact that both user data and applications are managed by the cloud provider. Clouds typically use a standardized data and application security architecture, while the demand for consumer and business services is steadily growing and becoming more diverse and with more sophisticated software design that leads to continuous upgrades to data security models in cloud computing. Due to the different methods of implementation in cloud computing service delivery models, the demand for a reasonable level of data protection is of utmost importance. The purpose of this publication is to propose a new information security model, which offers a solution for improving the use of sensitive data by introducing a three-factor authentication – an improvement of preventive control.*

Keywords: CLOUD SERVICES, INFORMATION SECURITY, CLOUD SERVICES ARCHITECTURE, CLOUD SECURITY, DATA SECURITY MODEL

1. Introduction

In the traditional calculation model, the data and software are saved entirely on the user's computer, while in the cloud model the user's computer may contain almost no software or data [1]. Cloud technologies are based on the idea that applications and software can be completely moved to an "invisible" place with shared resources on the Internet [2]. Cloud services offer a virtual platform of flexible shared resources (such as hardware, software and datasets), which are provided dynamically on demand with minimal effort made by users [3]. The main features of cloud services are the following [4]:

- Resource flexibility: Users and businesses can add or remove given resources on the basis of their immediate need. For example, at peak times, they can use more resources, which can be easily and quickly released during the rest of the time;
- Pay-as-you-go: Users pay only for the resources they use, as well as for the time they use them and for the other parameters;
- Multi-tenancy (Shared resources): Cloud services are based on a business model, whereby resources are shared at different levels - e.g. a network layer and an application layer. This means that numerous users use the same resources;
- Self-hiring of resources: Users can hire additional resources independently without much efforts (memory, software, network resources, etc.);
- Extremely scalable: Cloud technologies enable scaling up to tens of thousands of systems, as well as server space and memory.

To find a sufficiently suitable data security model in cloud technologies and especially models related to it is a particularly difficult task in today's ongoing changes in businesses and the IT sector.

Sh. Ajoudanian, for example, proposes a new data security model based on the CIA triad (Confidentiality, Integrity, Availability) and shows the role of the respective user and provider for each model [5]. However, the proposed model does not guarantee the security of sensitive data, as it allows simple user authentication. If you aim at a higher level of data security and protection, such as very sensitive data of an organization or entity, particular attention should be paid to authentication, which is not considered in the proposed model.

Another new data security and protection model is proposed by M. Mohamed, who considers three layers [6]: In the first layer, he proposes the authentication to be placed; in the second layer, he puts encryption of information, data integrity and protection for private users; and in the third layer, he proposes fast data recovery to be provided.

With the proposed model, the problem of sensitive data is partially resolved but the question remains as to whether it can be widely used. There is also a problem in terms of hybrid clouds, for example, if an organization has a private, public and community cloud, how will the three clouds be authenticated and how will the community and private clouds be protected, if the user logs in through the easily accessible public cloud and tries to gain unauthorized access to the other two hybrid clouds.

The Internet is not a place that cloud service providers can fully control and manage, which requires additional methods and techniques to ensure the security and protection of all participants in the cloud infrastructure.

The security of information and cloud infrastructure participants is of paramount importance to any cloud service provider and organization. As cloud technologies use a virtual environment, particular attention should be paid to the inherent risks, vulnerabilities and threats that threaten the security and protection of information and programs and differ significantly from the inherent risks, vulnerabilities and threats of a physical system.

This publication proposes a new data security model that improves the preventive control and the use of sensitive data by introducing three-factor authentication.

2. Security of service delivery models in cloud computing

Due to the specifics of each service delivery model in cloud technologies, the security is considered according to SLAs between the user and the provider for each of the following three main models:

1. Software as a Service

With this model, users pay a subscription for a software product, whereby part of the information or the full information is saved on a remote location and users can access this service via the Internet [7].

With SaaS, users do not have any control or authority to modify or manage cloud infrastructures or even specific applications that

are already developed [8]. Users have limited options to configure settings related to the use of these applications. The provider fully controls the cloud infrastructure and is responsible for the confidentiality, integrity and accessibility of the data and information.

2. Platform as a Service

This model enables users to develop their own cloud infrastructure applications using programming languages and additional software tools, which are provided by the cloud service provider (such as .NET, Ruby, Java, etc.). PaaS provides users with all the resources they need, so as to be able to develop applications and services entirely on the Internet without the need to download or install additional software. The user is not yet able to manage the underlying cloud infrastructure but only the applications developed by him/her [7].

One disadvantage of PaaS is the lack of interoperability and transfer of applications, which are already developed by the user, to other providers. I.e. if the user wants to transfer the applications developed with the current provider to another cloud provider, this cannot be done or if it can be done, the costs will be extremely high. Another disadvantage is that if the provider decides to leave the business, the applications and information in them will be deleted. With this model, the providers' liability is associated only with the integrity and accessibility of data. The user is responsible for the confidentiality and protection of the information.

3. IaaS

The infrastructure as a service allows the respective organization to create its own software environment. In the SaaS and PaaS models, the provider provides the user with applications but in the IaaS model, this is not done [7]. This model only provides the hardware on which the organization can install whatever it wants.

The control, exercised by providers, is at an extremely low level. They are only responsible for the accessibility of the services provided by them. Compared to providers, the users' responsibility is at a quite high level. They are responsible for the confidentiality, integrity of the data and its protection. The following table shows the responsibilities of cloud service providers and users for each model.

Before an updated version of a data security model in cloud technologies is proposed, the evolution of cloud technologies should be known very well, as well as their advantages and disadvantages. The following figure shows the overall development of these technologies by years.

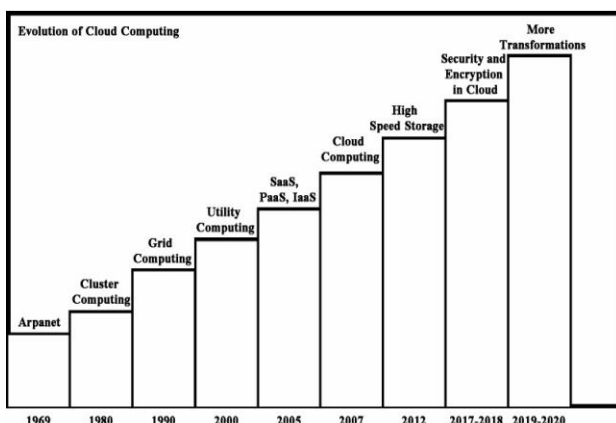


Figure 1. Evolution of Cloud Computing

The beginning of cloud technologies dates back to 2005, when the main models for the provision of cloud services were presented for the first time, namely the Software as a Service, Platform as a Service and Infrastructure as a Service. Subsequently, Cloud technologies began to focus more on data storage, accessibility and

other indicators. In 2012 the main goal was to increase the data and information storage speed (High Speed Storage).

In early 2016, cloud technologies encountered one of the biggest problems in this field, which have not yet been resolved - security and protection of information and data. Figure 1 also shows that the main work done in 2017 and 2018 focused on ensuring security and protection of these technologies.

In recent years, emphasis has been placed on creating more possibilities and transformations in cloud technologies, such as productivity, speed, security, etc. Unfortunately, the problem with the security and protection of these technologies remains unresolved.

In cloud service delivery models, it is important to properly allocate the responsibilities of the user and the provider for each model. Table 1 presents these responsibilities with respect to the so-called CIA triad.

Table 1. Responsibility of providers and users in cloud computing models

	Provider	User
Software as a Service	Confidentiality, Integrity and Availability	X
Platform as a Service	Integrity and Availability	Confidentiality
Infrastructure as a Service	Integrity	Confidentiality and Availability

As it can be seen in the table, for each model, the responsibilities of the user and the provider are different. For example, in case of the most widely used model globally - the Software as a Service - the provider is required to ensure the confidentiality, integrity and accessibility of data and information, while the user has no responsibilities. These are the so-called services or products for which a subscription fee is payable for the right to use the product or service. All new versions, patches, new modules, security and data protection, information, etc. are the sole responsibility of the provider of cloud services or products. In the other two models, the responsibility is divided between the provider and the user. For example, in case of the Infrastructure as a Service, the responsibility of the provider is only to ensure the integrity of the data and information (i.e. the provision of basic resources). Any other things are the responsibility of the user.

3. Information security in cloud computing

With the development of modern information technologies and systems, the problem of data security and protection remains one of the biggest problems of the 21st century. It is a particularly sensitive issue of cloud technologies and related service delivery models.

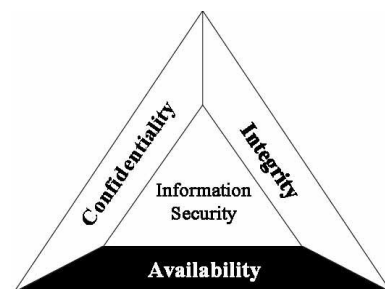


Figure 2. CIA triad of Information Security

To secure information in cloud technology, we should consider the main information security triad, the so-called CIA triad.

On the basis of this triad, all modern models of information security are considered.

1. Confidentiality of information in clouds

Information confidentiality aims to prevent disclosure of information by unauthorized users or systems. There are two categories of privacy in cloud infrastructure: privacy in a private cloud and privacy in a public cloud. Private cloud privacy is not a major concern, as it relates to private networks in an organization. There are several privacy issues with public clouds. The first one concerns controls of authentication and authorization. As numerous users use public clouds, they must be properly authenticated in the system by additional rigorous policies and procedures in addition to username and password controls. Some of the most up-to-date security models aim at either two-factor or three-factor authentication or web proxy logon [8] using Security Assertion Markup Language (SAML). With SAML, each organization manages its users by using trusted links to share authentication between sites. SAML is a very good solution for scaling up authentication. It uses the so-called Single sign-on, which is a process by which the user enters a username and a password to access numerous sites/applications with the same data securely and reliably [9].

2. Data Integrity in Clouds

Data integrity is the overall accuracy, completeness, and consistency of data. Data integrity also refers to the safety of data. Data integrity is not only whether the information is correct but whether it is reliable. In addition to this, users often use different cloud service providers and in some cases there is no guarantee of data security at the input and transaction management levels. Cloud providers must implement data integrity standards in their cloud services [9]. Some of the data integrity standards for today's cloud services are Data Integrity Field (DIF), SNIA Cloud Management Interface (CDMI), XML Based Solutions, etc.

3. Availability of Data in the Cloud

Data availability is the process of ensuring that data is available to end users and applications, when and where they need it. It defines the degree or extent to which data is readily usable along with the necessary IT and management procedures, tools and technologies required to enable, manage and continue to ensure data availability. One of the most significant security challenges for cloud technologies is the availability of data. Many providers aim at improving the availability of cloud infrastructures. There are several ways to improve the accessibility of cloud services. One of them is to provide users with back up information. A better option is to provide them with a caching proxy server that can respond to queries without contacting the specific server using content retained from a previous request by the respective user or even by another user. Another option is the transfer data from an online server to a so-called "hot" server that always has the contents of a cloned server and in case of a server failure, the user's information is not lost [10-13].

4. A new data security model in cloud computing

The development of the new data security model in cloud technologies and related service delivery models is based on the CIA triad for cloud services, which is described above. The CIA shows the responsibilities of providers and users for each of the three most widely used cloud service delivery models. For example, in the case of the most widely used model globally – the Software as a Service, privacy, data integrity and availability are the sole responsibility of the provider of cloud services or products. For the other two models – the Platform as a Service and Infrastructure as a Service, the responsibilities are allocated differently between the user and the provider, as shown in Table 1. In this allocation, the

most significant problem with cloud technology remains - the problem of data security and protection.

The new data security model aims at increasing the level of security of data and programs by checking users with the so-called three-factor authentication from another server. Three-factor authentication uses a combination of the following methods:

- **something the user knows:** it is a symbolic sequence that is known only to a given user and is unknown to others. When it is claimed with the user ID, it is taken as proof of possession of the identity. Typical examples of something known are passwords, PINs, private keys, etc. The most common and accessible means of authentication are passwords, which are a sequence of characters chosen by the system administrators or by the users themselves.
- **something the user owns:** it is a material object (card or other medium) on which additional information is usually recorded. In the process of authentication, the user claims the owned object by placing it in a device that can read its contents and confirm or reject the declared identity. Such techniques, for example, can help authenticate users using electronic signatures to share information with different institutions.
- **biometric data:** it represents unique data that can belong to only one individual. Fingerprints, iris, retina, voice recognition, etc. can be included in this category.

The model achieves a higher degree of security, as it focuses mainly on the preliminary control – the highest degree of efficiency. SSL or IPSec, as well as DoS protection, are added here.

The model also uses the so-called Single sign-on, which also requires authentication from an external server. Once the user is admitted to the system (cloud), he/she can access all cloud infrastructures in the environment (private, public, community) through Single sign-on without the need to enter authentication data again. That is, once a user is successfully admitted to the cloud infrastructure, this means that he or she has successfully passed the three-factor authentication described above and has the right to use his or her access to enter other cloud infrastructures in the environment, without the need to authenticate to each cloud individually (as it will be the case without a Single sign-on).

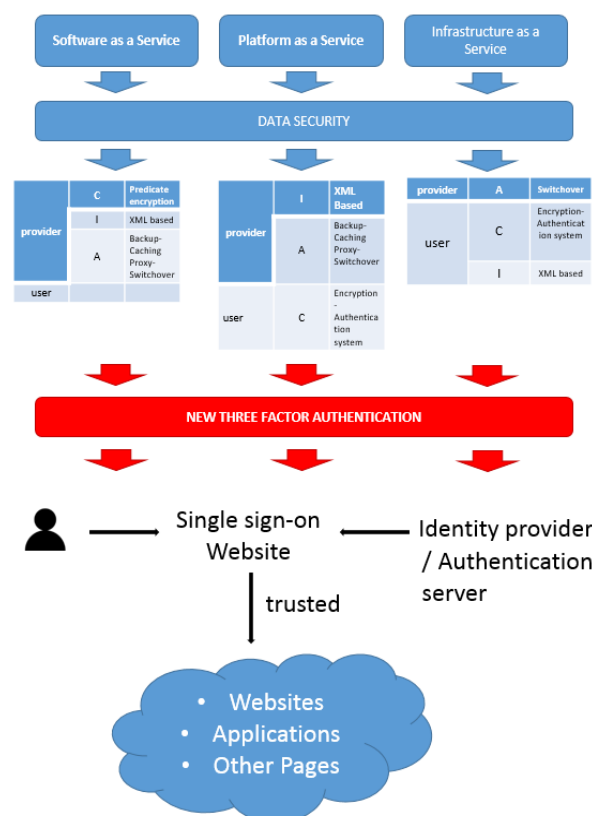


Figure 3. A New model for data security in cloud with three-factor authentication (based on CIA triad)

The model encrypts the information before it sends it to users by SSL, IPsec protocols. Another improvement of the model is that authenticated users and end users cannot be eavesdropped, which guarantees a high degree of confidentiality of the model. The model is mainly used in a cloud environment that requires maximum security and protection of data and information and enables the successful collection of biometric data.

Cloud providers are primary responsible to secure the information. Each provider uses special techniques to secure its resources in this system. Upon entry of a username and password, Single sign-on also allows some authenticated users to use this registration for other sites and applications.

5. Conclusion

After the massive implementation of cloud technologies in businesses and enterprises, the security of users' information and authentication has become increasingly important issue, which is discussed by many organizations around the world. Because of security issues, many users and businesses still refuse to use the cloud-based model. Risks, threats and vulnerabilities in virtual environments of cloud technologies significantly differ from the ones in physical environments. This publication introduces a new CIA triad data security model. It adds a three-factor authentication (3FA), as well as a Single sign-on using the OpenID standard. Thus, the effectiveness of preventive control is increased – this is the most important and effective control in an organization. The model focusses on data security and protection. Security management in an organization should also be in line with IT policies and standards for cloud technologies, as well as with business objectives for data security through the CIA triad – confidentiality, integrity and availability of information.

The proposed model will significantly increase data security in cloud technologies and thus reduce the risk of misuse and misuse of personal data, as well as the misuse of identity. Cyber threats and cyber-crimes will be reduced. The use of smart technologies will be safer and psychological stress and distrust in consumers will be reduced.

6. Acknowledgment

This work is supported by the Bulgarian National Science Fund, Project: KP-06-N 32/4/07.12.2019.

7. References

- [1] El-Gazzar R., Hustad E., (2016), Olsen D, Understanding cloud computing adoption issues: A Delphi study approach, Journal of Systems and Software, Volume 118, pp. 64-84. <https://doi.org/10.1016/j.jss.2016.04.061>
- [2] Viega J., (2009), Cloud Computing and the Common Man, in Computer, vol. 42, no. 8, pp. 106-108, DOI: 10.1109/MC.2009.252
- [3] National Institute of Standards and Technology(NIST), U.S. Department of Commerce, The NIST definition of Cloud Computing, Special publication of Computer Security, 2011
- [4] Center Of The Protection Of National Infrastructure CPNI by Deloitte, "Information Security Briefing of Cloud Computing"; The Security of Cloud Services, A Middle East Point of View, 2019
- [5] Ajoudanian Sh., Ahmadi M., (2012), A Novel Data Security Model for Cloud Computing, IJET 2012, Vol. 4 (3): pp. 326-329, DOI:10.7763/IJET.2012.V4.375
- [6] Mohamed M., Abdelkader H., El-Etriby S., (2012), Enhanced data security model for cloud computing, 2012 8th International Conference on Informatics and Systems (INFOS), Cairo, 2012, pp. CC-12-CC-17.
- [7] Hoyer U., Obel H., (2017), Guide on SaaS vs PaaS and IaaS, <https://www.linkedin.com/pulse/guide-saas-vs-paas-iaas-ulrik-hoyer-hansen-obel>
- [8] Zhiying W., Nianxin W., Xiang S., et al., (2020), An empirical study on business analytics affordances enhancing the management of cloud computing data security, IJIM, Volume: 50, Pages: 387-394, DOI: 10.1016/j.ijinfomgt.2019.09.002
- [9] Sheikh M, Liang J., Wang W., (2020), Security and Privacy in Vehicular Ad Hoc Network and Vehicle Cloud Computing: A Survey,

Wireless Communications & Mobile Computing, Volume: 2020, DOI: 10.1155/2020/5129620NIST, Cloud Computing Standards Roadmap, Special Publication 500-291

- [10] Garvanov I., H. Kabakchiev, V. Behar, M. Garvanova, R. Iyimbor, (2019), On the Modeling of Innovative Navigation Systems, Business Modeling and Software Design. BMSD 2019. Lecture Notes in Business Information Processing, vol 356. Springer, Cham, pp. 299-306, 2019. DOI: 10.1007/978-3-030-24854-3_23
- [11] Garvanova M., Shishkov B., (2019), Capturing human authority and responsibility by considering composite public values. Business Modeling and Software Design. BMSD 2019. Lecture Notes in Business Information Processing, vol. 356, 290-298. Springer, doi: https://doi.org/10.1007/978-3-030-24854-3_22.
- [12] Garvanova M., Shishkov B., Janssen M., (2018), Composite public values and software specifications. Business Modeling and Software Design. BMSD 2018. Lecture Notes in Business Information Processing, vol. 319, 412-420. Springer, doi: https://doi.org/10.1007/978-3-319-94214-8_32.
- [13] Garvanova M., Shishkov B., Vladimirov S., (2018), Mobile devices – effect on human health. Proceedings of the Seventh International Conference on Telecommunications and Remote Sensing – ICTRS'18, October 8-9, 2018, Barcelona, Spain, 101-104. New York: ACM, doi: <https://doi.org/10.1145/3278161.3278176>.

CAD/CAM-system module for the design of automatic production

Oleg Nikolaevich Mikhalev¹, Alexander Sergeevich Yanyushkin²
Faculty of Mechanical Engineering – Chuvash State University, Russia²

Mih_tm@mail.ru

Abstract: Automated production becomes a new reality of the industry, one of the important points in the construction of such production is a more accurate and quick design of technological processes for processing and assembly of parts. Analysis of computer-aided design systems shows the inadequacy of their capabilities to achieve this result. Significant improvement of CAD / CAM systems is required, for which it is necessary to establish close relationships between the part, tool, machine tool, technology, as well as digitize world experience and build an advanced algorithm that independently takes many different decisions, such as grouping part elements into operations and grouping transitions for their processing.

Keywords: CAD/CAM-SYSTEM, AUTOMATED PRODUCTION, AUTOMATION TECHNOLOGICAL PROCESS.

1. Introduction

A production site with an automatic production cycle of parts and the complete absence of people is the new reality of modern industry. All enterprises engaged in serial production go to this model of organization of production.

Automated production includes complex mechanical equipment with built-in electronic devices, such a combination creates electromechanical components that operate under the control of microcontrollers and a personal computer. These mechatronic modules have gained wide popularity in all fields, and by combining them whole mechatronic systems are created, for example, automatic production lines.

An automatic line may include CNC machines, robots, conveyors, manipulators, unmanned carts and many other devices. Since all components are interconnected through digital communication and controlled by a computer, they can work without people. This gives the most accurate, high-quality and high-performance result, excluding the human factor, which in turn leads to favorable economic indicators of production costs, and high competitiveness in the market.

Almost all automakers can serve as an example of such automated production; the car body moves along the line from one robot to another, which carry out welding and further assembly of the car without people.

Any production begins with its design, and the more thoroughly it is completed, the more efficient the production itself will be. Since automated production requires maximum debugging of all devices, the design of automated production is a more difficult task and is carried out using various CAD / CAM systems.

2. Design Automated Production

The scheme of automated production for each nomenclature of parts can be completely different, therefore its development is still a more creative process, however, production indicators, such as productivity, efficiency, profitability and others are more tangible and are expressed as specific numbers. The calculation of these numbers is very important and quite amenable to automation.

A lot of things determine the technological process of processing a part or a group of parts, since at the stage of development of the technological process the necessary metalworking equipment, devices, tools and many other equipment and accessories are selected, as well as processing strategies. From how rationally all the details of the technological process are thought out in many respects, they will determine the production indicators.

A lot of factors are involved in the development of the technological process that need to be taken into account; not taking into account the slightest action can lead to inconsistency and

downtime of expensive equipment, and as a result, to obtaining not maximum production indices.

Therefore, it is necessary to design the processing of each element of the part separately as efficiently as possible, accurately calculate all the processing parameters, and also form a sequence of their processing in each operation. This is a rather time-consuming process that requires a lot of knowledge, experience and time, and it is often necessary to optimize the process for various requirements, which takes additional time.

After the development of the technological process for processing the part, it is possible to develop the automated production scheme itself. At this stage, the composition of the equipment, their rational arrangement, the general sequence of moving parts play an important role, since a small inaccuracy in the location of only one element can lead to significant losses of time and other resources. To find the best production indicators, different schemes are compared, while the most accurate calculation gives a virtual simulation of the production process in real time, i.e. virtual modeling of different schemes.

An example of a computer-aided design system with precise detailing and visualization of all production processes is the Siemens Technomatrix system (see Fig. 1). Two-dimensional or three-dimensional virtual production is created in the system, object-oriented programming is implemented, with the help of which it is possible to create any non-standard processes, for example, moving a part by a person along a given trajectory at a given speed or assembling a node with hard-to-reach places by a robot. In the database of the system there are many elements of automated production and thus the organization of production with any degree of automation is possible.

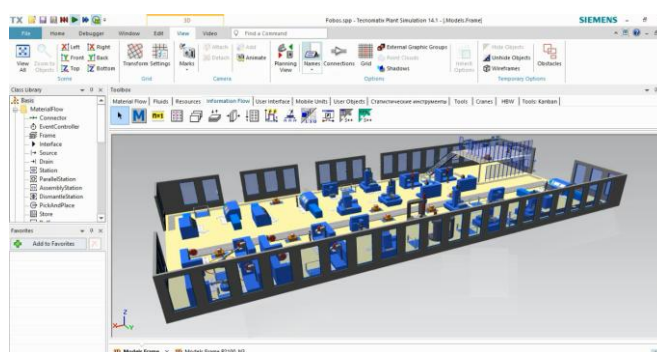


Fig. 1 Automated Production

So, setting the initial data on the technology of processing parts, processing time, you can get the most accurate performance, the designed production, as well as a single section or machine. By changing various elements, for example, moving equipment, replacing a person with a robot, as well as other optimizations, you can see how production indicators change. Thus, it is possible to achieve such a combination or arrangement of all elements of

automation of production, which will bring the maximum possible performance.

3. Automation Technological Process

The main trend in computer-aided design of technological processes for processing parts is the recognition of structural elements of parts, with automatic selection of the processing route and the necessary tools in accordance with a given logic or rules (see Fig. 2).

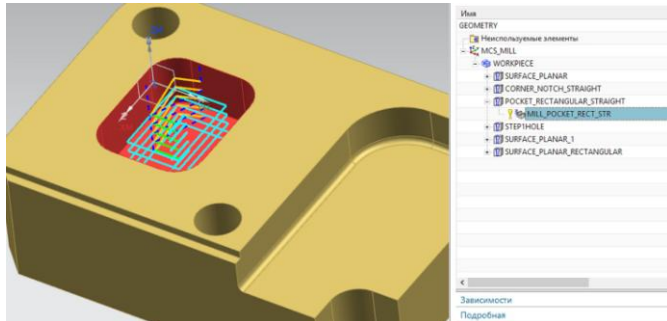


Fig. 2 Recognition of elements in the CAD/CAM-system

Automatic selection of the processing strategy with the selection of tools was implemented in the «CAD TPP CNC» for designing the processing of complex, precise holes, the processing of which requires more than 6 ... 10 transitions. Manually designing the processing of such holes takes a lot of time, this method made it possible to design the processing of complex holes in a matter of seconds using the one-button method, with all the necessary tools in the database [1].

Modern CAD / CAM systems implement similar capabilities for processing typical structural elements, taking into account the accuracy and roughness indices specified in the 3d model, as practice has shown, this significantly reduces the time of technological design. This possibility so far concerns only simple elements, and for non-typical and frequently encountered elements, it is possible to set individual logic.

However, the task of logic is often required for many typical elements, since the same element can be processed with different strategies, transitions and tools, it is necessary to teach systems to choose the most efficient processing method for a particular case, taking into account the characteristics of the machine, the processed material and other data.

To improve the method of recognition of elements, it is necessary to use more complex algorithms and a more developed database. Such improvement must be carried out taking into account modern requirements of production.

Recognition of elements is a big step in the automation of technological processes, but still insufficient, since modern automated production requires a higher degree of automation of design work. This requires not only recognition of the structural elements of the part, but also their grouping in operations, as well as the effective grouping of all transitions for processing grouped elements. Effective recognition in combination with a progressive grouping of operations will give a tremendous synergistic effect, which can significantly reduce the time and complexity of design.

Automating such a grouping of elements and technological transitions is a difficult task, since it is influenced by a huge number of factors. In addition to scientific research, account must be taken of practical tests and many other data. It is necessary to unite and consolidate world experience. Digitization of this experience will provide a developed database, for which it is necessary to create a complex algorithm that can make decisions and self-learn. Thus, it is possible to create automation of the design of technological

processes of a completely new level, the level of artificial intelligence.

The possibility of manual editing of the technology should be implemented very carefully, it should be as simple and quick as possible, for example, by highlighting a particular structural element in a 3D model and moving it to another operation, then the system itself will integrate the processing of the new element into the operation. And also the system should check all human actions for the possibility of this change, and exclude the human factor. Frequent the same manual changes to the technology, the system should fix and offer to adapt to them.

During the design of technological processes for automated production, a solution of many non-standard problems is required. For example, the selection of grippers for moving parts, the calculation of gripping forces, storage of parts, their correct positioning during movement, installation in the device, etc. For each detail, these tasks are solved individually.

Part of these tasks is difficult to automate, therefore it is necessary to create auxiliary modules that help technologists make the right decisions, as well as select or find new ideas for specific production conditions. Similar tasks must be solved in conjunction with a new approach to the automation of technological processes - recognition of structural elements with grouping of operations and transitions.

So the technologist will be able to quickly design a progressive technological process, select all the necessary devices for capturing and moving workpieces, calculate more accurate technological parameters for processing and movements, i.e. to get the exact necessary numbers with which you can create a scheme of automated production.

4. New Module of CAD/CAM-System for Automated Production

4.1. Technological Process

As has already been established in the automation of the design of technological processes in modern CAD / CAM systems, there are many bottlenecks, many tasks are solved manually or in a semi-automated mode, which is a laborious and inefficient design process.

The trend of automatic recognition of structural elements of a part plays an important role, but this is not enough, higher-order automation is necessary, namely, automatic grouping of operations and transitions. Also, many additional modules are needed to solve auxiliary problems of technological design.

In this connection, a new module of the CAD / CAM system is needed, which will link the solutions of these problems into a single whole. And also adapt the module in the CAD / CAM system, replacing existing recognition modules.

An analysis of the problems discussed above leads to the fact that in CAD / CAM systems there are no close links between the part, the processed material, the technology, the control program, the tool and the equipment (see Fig. 3).

Having established these close ties, as well as creating a developed database summarizing world experience, it becomes possible to automate such complex tasks as automatic selection of tools, machining strategies, calculation of cutting modes, grouping of transitions and many others, which will significantly reduce the time and complexity of design.

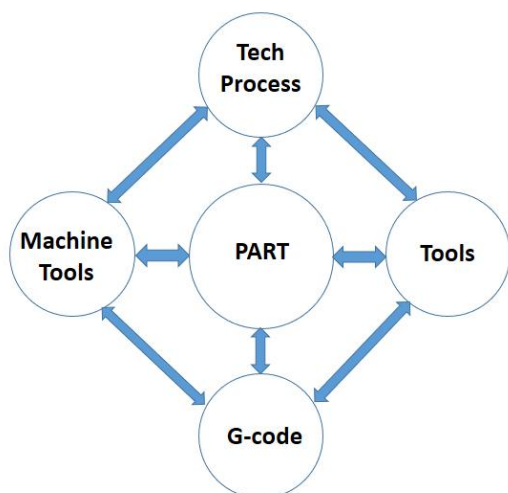


Fig. 3 The relationship of CNC-program with design elements

For example, setting the characteristics of the material to be processed, the rigidity of the part, machine tool, fixture, you can automatically select the strategy for processing the part and the corresponding tools, as well as calculate the cutting conditions, and then group transitions.

In this case, it is necessary to take into account various recognition and grouping filters, for example, recognition of elements located on certain sides of a part or grouping transitions with minimal tool changes or minimal part rotations, etc.

In the algorithm of the new module, it is necessary to take into account the rigidity of the technological system, the capabilities of the machine, otherwise the design result will be far from reality, and accordingly, performance indicators, including economic indicators, will be distorted.

In the «CAD TPP CNC», a relationship was realized between the parameters of each element of a complex hole and the methods of processing it, the processing routes, which in turn are associated with tools, and all this data is contextually related to each other [2].

This way of organizing data is also called a semantic model, i.e. different processing properties are inherited from each other, for example, when processing a hole, the processed material and equipment are already predetermined at the stage of tool selection; therefore, a drill compatible with this data is searched. So after recognizing the hole, the system, knowing about the machine, the material being processed, as well as other data, will independently select the necessary tools, and there is no need to set many additional filters. This also applies to the selection of processing routes and other technical data [3].

However, linking is not enough to group processing transitions for complex holes. To do this, it is necessary to give the system additional logic, for which a special algorithm has been developed that allows, based on the obtained interconnected data, as well as technological knowledge, to compose the necessary sequence of transitions for processing a complex hole of arbitrary shape.

Many complex holes can be present on one part, for the processing of which the same tools are used, in this case the system should group the processing transitions of the entire part in the most optimal way without unnecessary turns and tool changes.

4.2. Technical and economic indicators of automated production in the new CAD / CAM system module

According to the theory of productivity, technical production indicators are mathematically related to economic indicators. By

asking in this mathematical model technical data on technology, processing time, cutting conditions, productivity, composition of automated production, we can find out the cost of production, production costs, payback period, economic efficiency, profitability, etc. Thus, we can quickly draw a conclusion on the economic possibilities of building this production.

On the other hand, this theory works in the opposite direction, i.e. by setting the desired values of the cost of production, economic efficiency, it is possible to obtain technical indicators of automated production, such as the composition of equipment, degree of automation, number of jobs, design dates, etc.

Thus, using this mathematical model and the parameters of the new CAD / CAM system module, you can quickly work out several options for automated production and find the most optimal one for a particular enterprise. For example, for one plant, production will be 100% automated using the most productive equipment, for the second 70% with partial use of human labor, and for the third it will be necessary to develop special multi-spindle equipment, etc.[4]

To study this mathematical model, it is necessary to describe the elements of production: types of automated equipment, auxiliary equipment, drives, transport, warehouse, etc., including establishing links between them, as well as criteria for their effectiveness. In addition, it is necessary to establish links between typical technological processes, routes and elements of automated production [5].

The construction of this mathematical model is a difficult task, because it requires a lot of different studies, in addition, it is necessary to take into account a number of parameters that are difficult to automate, for example, the parameter of production flexibility. Production can be fully automated, but it may be absolutely not adapted to change the manufactured parts. And in the world with a frequent change of products, the task of automating production requires taking into account the requirements of flexibility.

4.3. Digitalization of production

The passing Fourth Industrial Revolution aims to combine industrial and digital technologies, during which the construction of virtual production, digital doubles, robotic production and many other changes.

These opportunities show rapid development and already today demonstrate a positive effect, and will soon become the main criteria for the competitiveness of enterprises.

The introduction of various technical equipment is now tested in a virtual environment and only after that is their production and implementation carried out. Many physical tests are not required, which significantly reduces the cost and speeds up the process of introducing new equipment into production.

When simulating, several hundred calculations are carried out that determine the scheme and composition of automated production equipment, the manual method of carrying out these calculations takes a lot of time, which is unacceptable in modern conditions.

The new CAD / CAM system module is designed not only to design the most advanced technological processes in the most automated way, but also to build effective automated production of parts on their basis, determine the number of equipment, the number of jobs, and find the optimal technical and economic performance indicators that can be used to virtually real-time test production. If parts are changed, you can also quickly adapt their processing in this production.

Digitalization involves the conversion of all information into a digital form, thus creating databases, visualizing processes, automating calculations, developing high-order algorithms, introducing industrial networks, common industrial sites, digital

services and many other features. All this creates new requirements for the construction of automated systems.

For example, as the initial information, it is not the drawing of the part, but the 3d model, thus new requirements for the construction of the design 3d model appear, in which all the data indicated in the drawing should be reflected. According to this model, it is possible to carry out automatic recognition of elements and solve many other problems in an automated mode.

In this regard, in some cases, the design of drawings and technological processes becomes an unnecessary task, in the presence of a detailed 3d-model, or rather a digital double of the part, the designed technologies and control programs for CNC machines, documentation is no longer necessary, all these data are transmitted to the machine in paperless form. Thus, a number of enterprises today have already refused to issue design and technological documentation. However, for enterprises where complex products are manufactured, technological processes are still required, but if detailed digital data are available, generating it is a possible automation task.

4.4. Design of mechatronic modules

In the construction of automated production, the share of electronics involvement is steadily increasing, there is a transition from mechanics to mechatronics. In fact, mechatronics unites many different areas: mechanics, electrical engineering, microelectronics, information technology, power electronics and many others. As a result of this association, a new technique is created that implements any actions, for example, moving parts between machines. It is impossible to create such a technique on the basis of only one field of science; therefore, mechatronics products are very complex and are implemented by many specialists.

The mechatronic module is an independent product for the implementation of movements, and the mechatronic system is a collection of mechatronic modules that are synergistically connected to each other to perform a specific functional task, for example, an automated part production line.

The design of such mechatronic modules and systems is the team work of many specialists in mechanics, electronics, programming and many others - this creates the requirement of a single space in which all specialists will be interconnected when working on one project.

In the new module, we will consider only the mechanical part of the design process of automated production, but with the possibility of integrating modules for other specialists. Therefore, in the framework of the new module, it is necessary to solve the following problems: a more detailed selection of gripper and other manipulator nodes, and calculations for loading equipment. With a detailed drawing of all the components of the mechatronic modules, it becomes possible to more accurately calculate whether they will cope with the flow or not, based on the designed technological process.

Thus, at the design stage, it is possible to track that one machine or robot cannot cope with the flow, thus a number of solutions are applied to eliminate bottlenecks at the design stage. It is possible to add a robot, and it is possible to move the robot on its guides, or to take other measures.

4.5. Implementation of a new CAD / CAM system module for automated production

Summing up all of the above, for the design of automated production it is necessary: development of an optimal technological process, calculation of technical and economic indicators, digital representation and simulation, design of mechatronic modules, creation of a single space, as well as the solution of many related tasks.

Automation of a very large number of tasks, but all of them must be solved interconnected, supplementing each other with data, with a possible return to the initial data for their correction and new tests. This requires the creation of a large software module for the CAD / CAM system.

However, in the framework of the study on the creation of such a module, it is important to understand the relationship between various design data both within a separate task, for example, the tasks of computer-aided design of technological processes, and between data of different tasks, for example, data on the technological process and determining the composition of the equipment of an automated complex. Between the technological process and the calculation of technical and economic indicators.

The work of such relationships has confirmed its relevance in conjunction with the algorithm for grouping technological transitions for processing complex holes. This makes it possible to implement a higher order algorithm, namely the grouping of structural elements of a part in an operation, as well as the construction of automated complexes for processing these parts [6].

Such algorithms will endow the CAD / CAM system with the property of independent decision-making, which people today still decide.

5. Conclusion

The most important for the construction of automated production is the establishment of an effective technological process with the accurate calculation of all technological parameters: productivity, efficiency, etc. For this, a new module of CAD / CAM systems is needed, which allows solving this problem more automatically.

The establishment of close relationships between the part, tools, machine tool makes it possible to create high-order algorithms to achieve this goal, as well as the design of automated production.

In the context of the rapid development of digital technologies, automation of production, this task becomes one of the most important for the quick launch of new products on the market and ensuring its low and competitive cost.

6. References

1. Mikhalev ON, Yanyushkin A. S. Increasing the degree of automation of CAD / CAM systems in the design of precision hole machining on multi-purpose CNC machines / O.N. Mikhalev, A.S. Yanyushkin // Moscow, Bulletin of computer and information technology. - 2008. - No. 5. - S. 33.-38.
2. Bazrov B.M. Modular technology in mechanical engineering./ B.M. Bazrov - Moscow: Engineering, 2001. -- 368 p.
3. Yanyushkin A.S., Mikhalev O.N. The computer-aided design system for technological preparation of production on CNC machines (CAD TPP CNC v. 1.0) // Certificate No. 20088615315 2008.
4. Kuznetsov Y., Gaidaienko Iu. Genetic Modeling and structural synthesis of CNC multi-spindle automatic machines of new generation./ International Scientific Journal «Industry 4.0»/ Vol. 3 (2018), Issue 3, pp 115-119.
5. Design of automated machines and complexes T2: textbook: in 2 volumes / ed. P.M. Chernyavsky. - 2nd ed., Rev. - M.: Publishing House of MSTU named after N.E. Bauman, 2014.
6. Mikhalev O.N., Yanyushkin A.S., Intelligent processing module CAM-systems. Obrabotka metallov (tekhnologiya, oborudovanie, instrumenty) = Metal Working and Material Science, 2019. no. 3 (76), pp. 30–50. doi: 10.17212/1994-6309-2017-3-30-50. (in Russian).

Abstract: The peculiarities of carrying out the basic methods of solving the problems of identification, which have become widespread in transport technologies, are considered. Three approaches were tested: a deterministic approach based on the method of selected points, a statistical approach that includes the least-squares method, and - an inductive approach - a method of group accounting of arguments. A number of experiments have been conducted, which is based on structural optimization of the model. In the process of analysis of the obtained models, a rational model of setting and solving identification problems in transport technologies of optimal complexity was obtained. This systematic flow diagram can be successfully applied to solve complex analytical tasks.

1. Introduction.

Engineering approach

Scientific approach

Thus, the right choice of the optimal method of solving the problems of identification is very relevant and necessary, especially in modern transport technologies.

The purpose of the study is to identify a control entity that is an element of transportation technology and is influenced by several input quantities.

2. The methodology of scientific research.

The deterministic approach – the method of selected points [9] is one of the simplest methods of determining a formal expression. For example, the static characteristics of the system is the method of selected points. This method corresponds to a Lagrange or Newton polynomial. Static characteristics of the system can be set as a curve of any order, such as 2nd order:

but if the attempt is unsuccessful then the order of the equation must be increased. Yes, entering new variables:

we obtain the equation of the static characteristic in a linear form:

The algorithm for solving the problem of identification using the method of selected points is shown in Fig. 1.

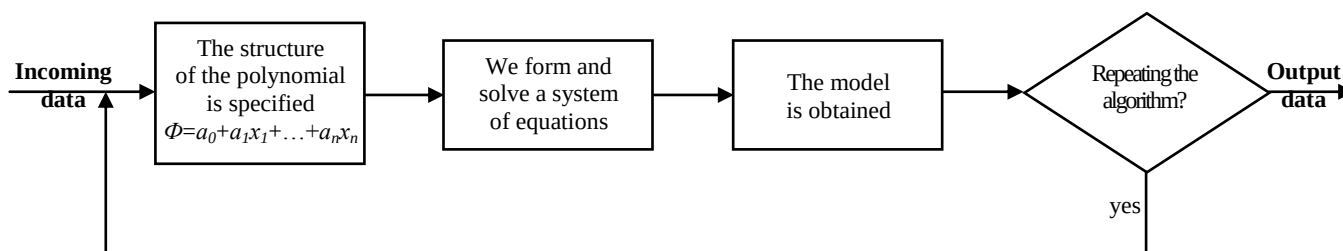


Fig. 1 Structural diagram of the method of selected points

Substituting a value to a given equation, we obtain a system of equations (4):

$$\left. \begin{aligned} &a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ &a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ &\dots\dots\dots \\ &a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{aligned} \right\}, \quad (4)$$

or in matrix form: $Ax = b$,

where $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$ – is the matrix of coefficients of the

system; $b = [b_1, b_2, \dots, b_n]$ – the vector of free terms; $x = [x_1, x_2, \dots, x_n]$ – vector of unknowns.

The numerical solution of the SLAR by the Gauss method is based on the reduction of the matrix of coefficients of system A to a triangular form, which is achieved by sequential removal of unknown equations from the system. Consider the scheme of a single division. Initially, the first equation removes the unknown x_1 from all the following equations. Why is the first equation divided by the principal element a_{11} (provided that $a_{11} \neq 0$), and then the equation multiplied by the corresponding coefficient at $x_1 - a_{i1}$ is subtracted from the last equations of the system. An intermediate matrix $A^{(1)}$ is obtained. Then the second equation removes the unknown x_2 , the intermediate matrix $A^{(2)}$ is obtained, and so on, until only one unknown x_n remains in the last equation. The equivalent system is thus obtained:

$$\left. \begin{aligned} x_1 + c_{12}x_2 + \dots + c_{1n}x_n &= d_1, \\ x_2 + \dots + c_{2n}x_n &= d_2, \\ &\dots\dots\dots \\ x_n &= d_n \end{aligned} \right\}, \quad (5)$$

unleashing which one will get the values of the variables that will characterize the model.

The method of least squares allows to determine the parameters of models of static and dynamic systems, based on the criterion of the minimum standard error.

The static system is defined by the system of functions of many variables, which are represented in the form of power expansions up to 9th degree, including:

$$f_i(x) = f(x_1, x_2, \dots, x_i), i = 1, \dots, n. \quad (6)$$

The statistical characteristic of a system with one perturbing effect and one regulating effect can be represented as a second-order curve:

$$\Phi = a_0 + a_1\lambda + a_2\mu + a_3\lambda^2 + a_4\mu^2 + a_5\lambda\mu. \quad (7)$$

If the sum of the squares of the deviations of the experimental values from the calculated values is greater than some predetermined number, then the order of the equation must be increased. The problem of determining the coefficients of the equation is a problem of parametric optimization [10]. If it is necessary to determine the form of an equation, then this identification is structural. Equation coefficients are determined using the least squares method, ie by minimizing the function:

$$\Omega = \sum_{i=1}^l (\Phi_i^{theor} - \Phi_i^{exp})^2, \quad (8)$$

where Φ_i^{theor} – calculated values of the function Ω , Φ_i^{exp} – experimental (a priori) values.

The “IDENT” program [11, 12], which is a part of the “FACTOR +” PMC, finds the best RMS approximation of the function of several variables of a given polynomial. The least-squares problem we consider here is called differently in different scientific disciplines. For example, mathematicians may approach it as a search task for a given point of the functional space of the closest point in a given subspace. Statistics introduce probabilistic distributions in their formulation and use terms such as regression analysis to describe this area. Engineers come to this task with subjects such as parameter estimation or filtering.

The main point is that when these tasks (formulated in any of these contexts) reach the stage of specific calculations, they have the same central problem, namely the sequence of linear least-squares problems.

This basic linear least-squares problem can be formulated as follows: let the data be a valid $m \times n$ -matrix A of rank $k < \min(m, n)$ and a real m -vector b . Find the true n -vector x_0 that minimizes the Euclidean length of the vector $Ax - b$.

Consider as an important example the case where a linear least-squares problem originates from a nonlinear problem and the solution vector should be used as an amendment to the current nominal solution of the nonlinear problem. A linearized task will be a useful approximation to a nonlinear one in only a limited neighborhood. If there are different vectors, a sufficiently small non-inferiority giving in a linear problem, then one may prefer that it has the smallest length. This increases the likelihood of staying in a neighborhood with a reasonable linear approximation.

Group Method of Data Handling (GMDH) [13] is a method of generating and selecting regression models of optimal complexity, the block diagram of which is shown in Fig. 2.

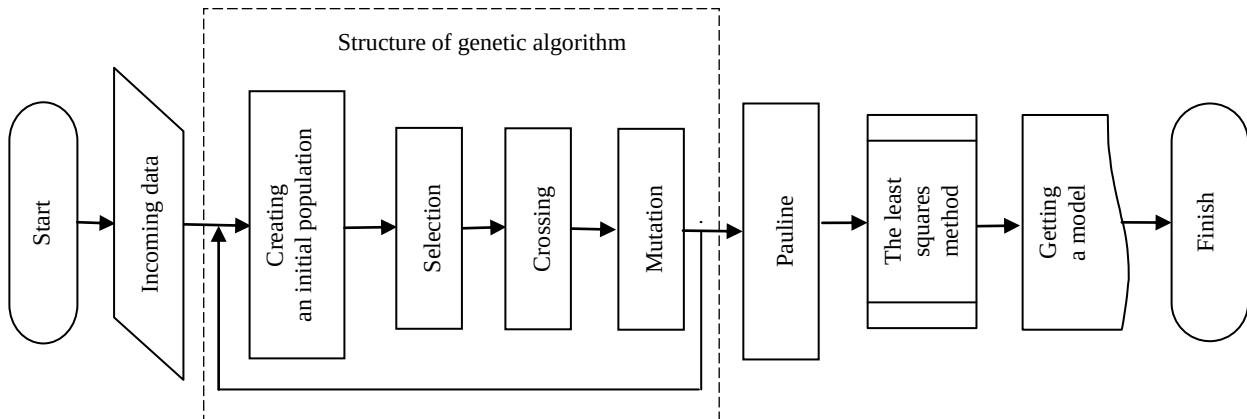


Fig. 2 Flowchart of the GMDH algorithm

The complexity of the model in GMDH means the number of parameters. To generate, we use a basic model whose subset of elements must be included in the desired model. External criteria, special model quality functionals calculated on the test sample, are used to select models.

GMDH is recommended for use when the sample contains several items. Then, when constructing regression models, it is impossible to use statistical hypotheses about the density of distribution, such as the Gaussian distribution hypothesis. Therefore, an inductive approach is used to consistently generate models of increasing complexity until at least some quality criterion of the model is found. This quality criterion is referred to as the external criterion because different data is used when setting up models and evaluating model quality. Achieving a global minimum of the external criterion when generating models means that a model that has reached such a minimum is in demand.

3. Software experimental verification of data.

The study of the optimization problem was carried out using the program "IDENT". Consider the operation of the program in two-dimensional space, that is, in the Cartesian coordinate system. Table 1 provides data for experiments with a single-variable function.

Table 1: Input data sampling

x	0,1	0,3	0,5	1	3	4	5
y	0,2	0,6	1	2	6	8	10

Table 2 shows the results of the program for four variants of polynomial structures that are consistently complicated.

Table 2: The results of the study of sequential complication of the polynomial structure (function of one variable)

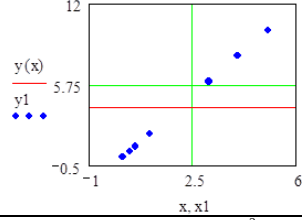
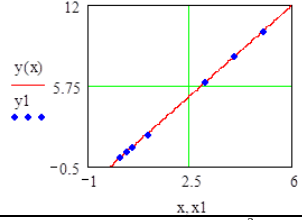
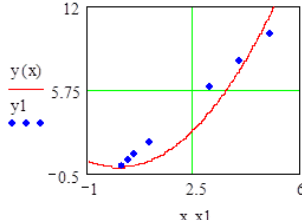
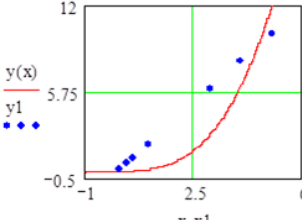
Experiment 1. $y = C_0$	Experiment 2. $y = C_0 x$
Polynomial coefficients: $c(0)=3.971429$ Error: error=3.7 The polynomial value in points: $ff(0)=3.971429$ $ff(1)=3.971429$ $ff(2)=3.971429$ $ff(3)=3.971429$ $ff(4)=3.971429$ $ff(5)=3.971429$ $ff(6)=3.971429$ Obtained polynomial: $y = 3.971429$	Polynomial coefficients: $c(0)=2.000000$ Error: error=0 The polynomial value in points: $ff(0)=0.200000$ $ff(1)=0.600000$ $ff(2)=1.000000$ $ff(3)=2.000000$ $ff(4)=6.000000$ $ff(5)=8.000000$ $ff(6)=10.000000$ Obtained polynomial: $y = 2x$
Experiment 3. $y = C_0 x^2$	Experiment 4. $y = C_0 x^3$
Polynomial coefficients: $c(0)=0.450960$ Error: error=1.2 The polynomial value in points: $ff(0)=0.004510$ $ff(1)=0.040586$ $ff(2)=0.112740$ $ff(3)=0.450960$ $ff(4)=4.058637$ $ff(5)=7.215354$ $ff(6)=11.273991$ Obtained polynomial: $y = 0.450960x^2$	Polynomial coefficients: $c(0)=0.094183$ Error: error=1.9 The polynomial value in points: $ff(0)=0.000094$ $ff(1)=0.002543$ $ff(2)=0.011773$ $ff(3)=0.094183$ $ff(4)=2.542945$ $ff(5)=6.027722$ $ff(6)=11.772895$ Obtained polynomial: $y = 0.094183x^3$
	
	

Table 3 shows the data for the two-variable function experiments.

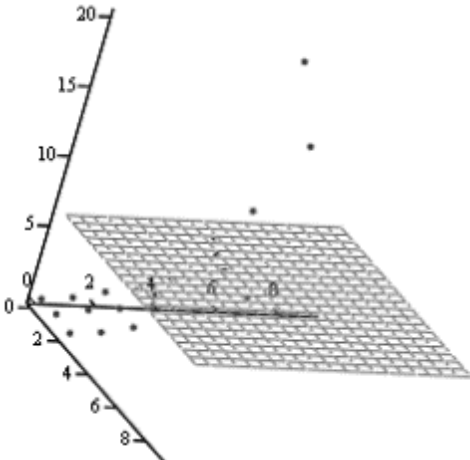
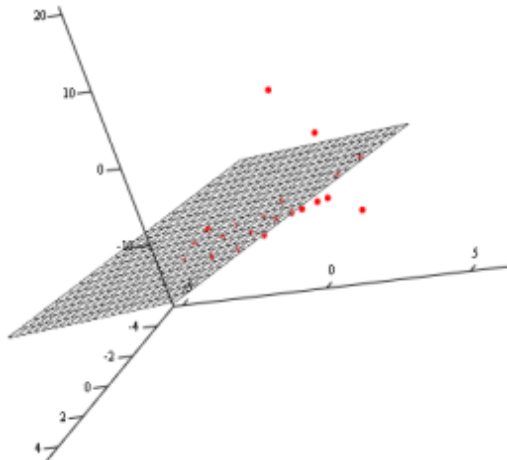
Table 3: Input data sampling

x_1	0,1	0,3	0,8	1	2	5	7
x_2	0,2	0,6	0,8	2	3	4	2
y	0,02	0,18	0,64	2	6	20	14

Table 4 shows the program results for five consecutively complicated polynomial structure variants.

In Fig.3 a) it is shown how, when complicating a function from one variable in the second experiment, the polynomial that most closely corresponds to a given set of points was found, with further complication of the polynomial the error started to increase, which led to the deterioration of the result. The graph of Fig.3 b) for a function with two variables. For this case, the polynomial with the smallest magnitude error was found in the third experiment.

Table 4: The results of the study of sequential complication of the polynomial structure (a function of two variables)

Experiment 1. $y = C_0$	Experiment 2. $y = C_1 x_1 + C_2 x_2$
Polynomial coefficients: $c(0)=6.120000$ Error: error=7.3 The polynomial value in points: $ff(0)=6.120000$ $ff(1)=6.120000$ $ff(2)=6.120000$ $ff(3)=6.120000$ $ff(4)=6.120000$ $ff(5)=6.120000$ $ff(6)=6.120000$ Obtained polynomial: $y = 3.971429$	Polynomial coefficients: $c(0)=1.865445$ $c(1)=1.489670$ Error: error=2.5 The polynomial value in points: $ff(0)=0.484478$ $ff(1)=1.453435$ $ff(2)=2.684091$ $ff(3)=4.844784$ $ff(4)=8.199898$ $ff(5)=15.285901$ $ff(6)=16.037452$ Obtained polynomial: $y = 1.865445x_1 + 1.48967x_2$
	

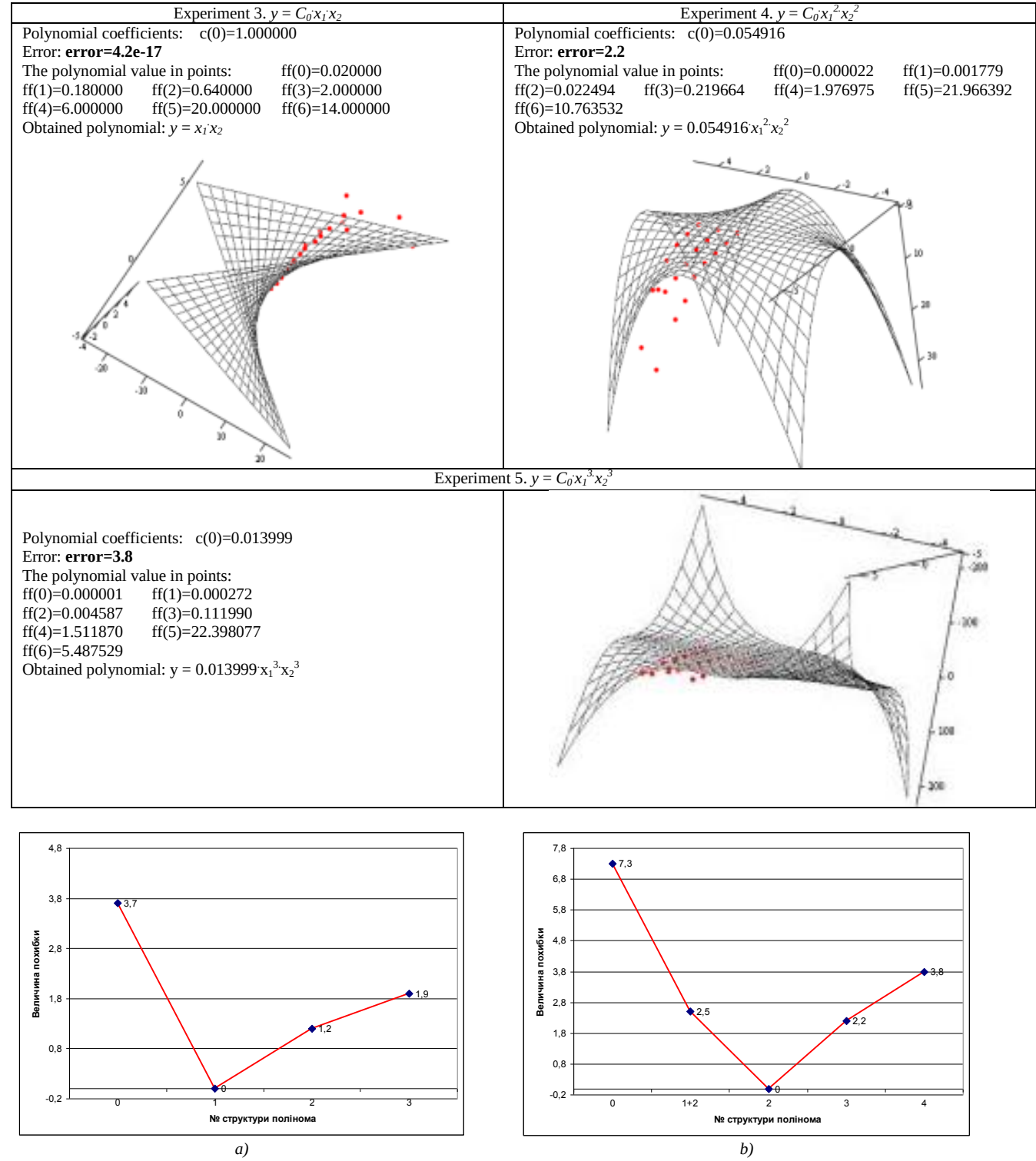


Fig. 3. Functional dependence of change of magnitude of error at complicated structure of a polynomial:
a) for one variable; b) for two variables.

Fig. 4 depicts a generic block diagram showing the interaction of the identified identification tasks. Structural optimization depicted in the graph fully incorporates parametric optimization, ie

the least-squares method, and also shows the functional interaction of structural optimization with the group argument method.

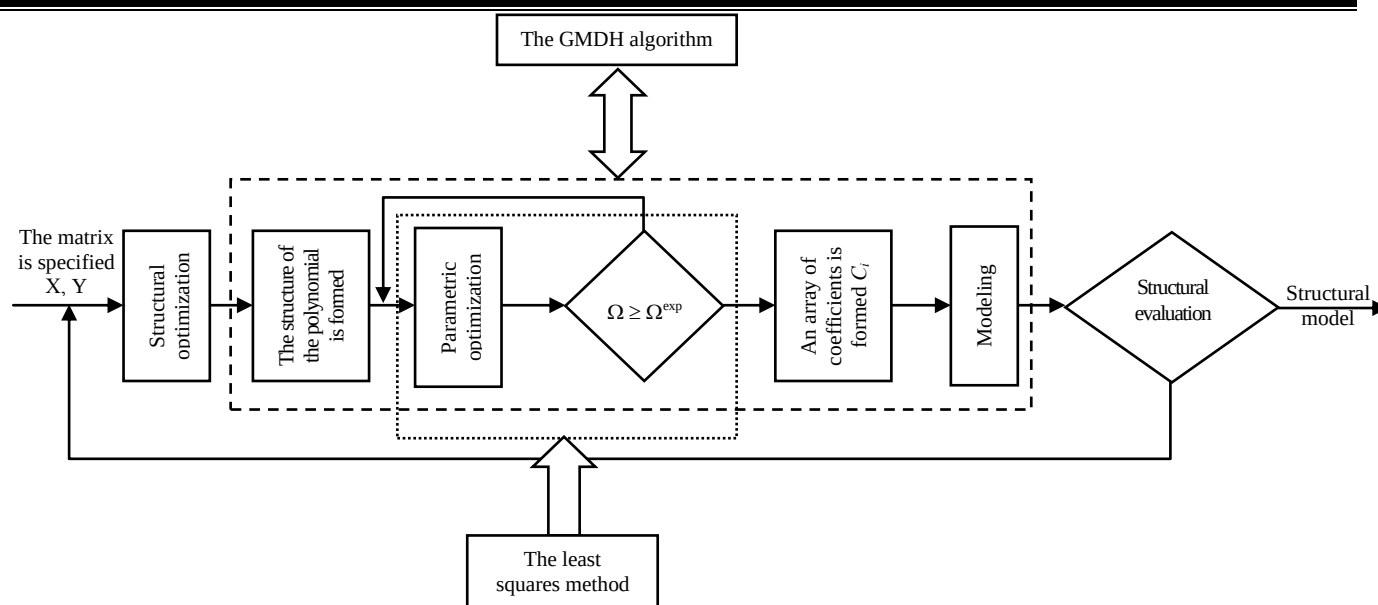


Fig. 4 Flowchart of the interaction of identification methods

4. Conclusion

As a result of scientific researches the peculiarities of realization of the basic methods of solving the problems of identification, which have become widespread in transport technologies, are considered. To do this, we analyzed and compared three methods for solving identification problems: a deterministic approach based on the method of selected points, a statistical approach that includes the least squares method, and inductive approach – a method of group accounting of arguments.

Based on the considered methods, an extended flowchart is constructed, which gives a generalized characteristic of the interaction of the described tasks. As a central element of the flowchart is structural optimization, on the basis of which experimental studies were conducted on the complexity of the polynomial structure and finding the model with the smallest amount of error.

A number of experiments have been conducted, based on structural optimization of the model. In the process of analysis of the obtained models, a rational model of setting and solving identification problems in transport technologies of optimal complexity was obtained. This systematic flow diagram can be successfully applied to solve complex analytical problems in transportation technologies.

5. Literature

1. S. Matthew, *Using computers in the law office* (Clayton State University, Clayton, USA, 2015).
2. F. Necati Catbas, A.E. Aktan, Tracy Lynn Kijewski-Correa, *Structural Identification of Constructed Systems: Approaches, Methods, and Technologies* (American Society of Civil Engineers, San-Francisco, USA, 2013).
3. Zh.Yafei, L.K.Keung, Sh.Wang, *Forecasting air travel demand* (Routledge, New-York, USA, 2018).
4. Spyros G. Tzafestas, *Systems, Cybernetics, Control and Automation* (River Publishers, Gistrup, Denmark, 2017).
5. P. Aileen, B. Niamh, *Principles and practice of group accounts* (Thomson, London, UK, 2003).
6. A.G. Ivakhnenko, V.G. Lapa, *Cybernetics and forecasting techniques* (American Elsevier Pub. Co., Wisconsin, USA, 1967).
7. P. Melin, O. Castillo, *Modelling, simulation and control of non-linear dynamical systems* (Taylor & Francis, London, UK, 2002).
8. T. Turpin, V.V. Krishna, *Science, Technology Policy and the Diffusion of Knowledge* (Edward Elgar Publishing, Cheltenham, UK, 2007).
9. A. Tymchenko, E. Kryganiivskii, V. Melnik, M. Pidgornii, *Bulletin of ChSTU*, **3** (2014) (in Ukrainian).
10. K. Stoilova, T. Stoilov, V. Ivanov, *Cybernetics and Information Technologies*, **2**, 17 (2017), doi:10.1515/cait-2017-0019.
11. I. Garvanov, C. Kabakchiev, B. Shishkov, and M. Garvanova, in: *Seventh International Conference on Telecommunications and Remote Sensing "ICTRS 2018"* (Barcelona, 2018), pp. 95-98, doi: 10.1145/3278161.3278175.
12. M. Garvanova, I. Garvanov, S. Vladimirov, *Industry 4.0*, **4**, 6 (2019).
13. A.A. Tymchenko, M.V. Pidgorniy, G.O. Zaspas, V.V. Boiko, *Bulletin of ChSTU*, **3** (2009).

A survey on deep learning in big data analytics

Makrufa Hajirahimova¹, Aybeniz Aliyeva²

Institute of Information Technology – Azerbaijan National Academy of Sciences, Baku, Azerbaijan
hmakrufa@gmail.com¹, aliyevaliyeva.a.s@mail.ru²

Abstract: Over the last few years, Deep learning has begun to play an important role in analytics solutions of big data. Deep learning is one of the most active research fields in machine learning community. It has gained unprecedented achievements in fields such as computer vision, natural language processing and speech recognition. The ability of deep learning to extract high-level complex abstractions and data examples, especially unsupervised data from large volume data, makes it attractive a valuable tool for big data analytics. In this paper, we review the deep learning architectures which can be used for big data processing. Next, we focus on the analysis and discussions about the challenges and possible solutions of deep learning for big data analytics. Finally, have been outlined several open issues and research trends.

Keywords: BIG DATA, BIG DATA ANALYTICS, DEEP LEARNING, AUTO-ENCODERS, DEEP BELIEF NETWORKS, CONVOLUTIONAL NEURAL NETWORKS, RECURRENT NEURAL NETWORKS, RESTRICTED BOLTZMANN MACHINES

1. Introduction

Recently, Deep learning and Big data analytics are the very active fields of research in the science and engineering domains. Big data is defined as digital data, which is difficult or impossible to manage and analyze with traditional software tools and technologies [1]. Analyzing of data and obtaining knowledge and useful information from them is very important for making motivated decisions in organizations, new scientific revelations, national security and healthcare fields. The demand for data analysis in real-time has led to the creation of Big data analytics. Big data analytics is a process of extracting useful information from large volumes of data to make optimal (best) decisions. The size of data has considerably grown in the last decade, with the emergence of social networks, Internet of Things, cloud computing and other technologies. The rapid increasing of data volume, along with the promises potential opportunities for all sectors of society, creates problems for data mining and information processing [2]. Dealing with these data can be supported by Deep learning capabilities, especially its ability to deal with both the labeled and unlabeled data which are often collected abundantly in Big data. Deep learning is an attractive research topics that belong in Artificial Intelligence (AI). DL refers to machine learning techniques that based on supervised and unsupervised methods to automatically learn hierarchical representations in deep architectures. It has achieved unprecedented success in applications of essential fields such as computer vision, speech and audio processing, and natural language processing [3-7].

The ability of Deep learning to extract high-level, complex abstractions and data representations from large volumes of data, especially unsupervised data, makes it attractive as a valuable tool for Big data analytics [4-6]. More specifically, Big data analytics problems such as semantic indexing, data tagging, fast information retrieval, and discriminative modeling can be better addressed with the aid of Deep Learning. In addition, there are need to use of Deep learning methods in solving of different problems that faced Big data analytics such as fast moving streaming data, highly distributed input sources, noisy and poor quality data, high dimensionality, scalability of algorithms, unsupervised and un-categorized data, limited supervised / labeled data and format variations of raw data.

The structure of the paper is organized as follows: Section 2 presents a brief review of typical deep learning models which are the most widely used for big data analysis and feature learning. Section 3 introduces possible solutions of deep learning for big data analytics challenges. Section 3 gives some open issues and research trends; and the final section is conclusions.

2. Brief review of deep learning architectures

Deep learning refers to a set of machine learning techniques that learn multiple levels of representations in deep architectures. In the last few years, various deep learning architectures have been developed. A brief overview of the deep learning architectures has been looked throw that are commonly used in Big Data analytics below.

2.1. Autoencoder and Stacked Autoencoders (SAEs)

As one of the most widely used deep learning techniques, Stacked autoencoders (SAEs) are constructed by stacking several autoencoders that are the most typical feed-forward neural networks [7]. Autoencoder is a kind of unsupervised learning structure that owns three layers: input layer, hidden layer, and output layer (Fig 1.). The process of an autoencoder training consists of two stages, i.e., encoding stage and decoding stage. Encoder is used for mapping the input data into hidden representation, and decoder is referred to reconstructing input data from the hidden representation.

SAE is typically trained by two stages, i.e., pre-training and fine-tuning. In the pre-training stage, each auto-encoder model is trained in a unsupervised layer-wise manner from bottom to top. This operation is repeated until the parameters of all the hidden layers are trained. After all the hidden layers are trained, backpropagation algorithm is used to minimize the cost function and update the weights with labeled training set to achieve fine-tuning [7, 8].

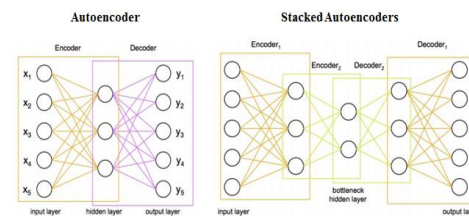


Fig. 1 Architecture of autoencoders.

2.2. Restricted Boltzmann Machines (RBMs) and Deep Belief Network (DBN)

Deep belief network is the commonly used, and successfully trained architectures in deep learning [9]. DBN is stacked by several restricted Boltzmann machines, as presented in Figure 3. RBMs are the most popular version of Boltzmann machine [5,7]. The Restricted Boltzmann Machine (RBM) is a probabilistic graphical model or a type of stochastic neural network. The network consists of two layers, i.e., visible layer and hidden layer (Fig. 2). The restriction is that there is no interaction between the units of the same layer and the connections are solely between units from different layers.

Deep belief networks have the potential to learn the representation of features using structured and unstructured data. It consists of input, hidden, and output layer. RBM uses DBN to construct a model that consists of two layers that are fully connected to each other. DBN combined strategies of unsupervised pre-training and supervised fine-tuning. The unsupervised stages intend to learn data distributions without using label information and supervised stages perform local search for fine tuning[7]. In the literature, DBN model is used by many researchers to efficiently and accurately process big data.

In particular, a graphical processing unit (GPU)-based model using stacked RBM in parallel to handle large volume of data with minimized process time. The power of deep learning is that it can train and handle millions of parameters at a time. Several restricted Boltzmann machines can be stacked into a deep belief network.

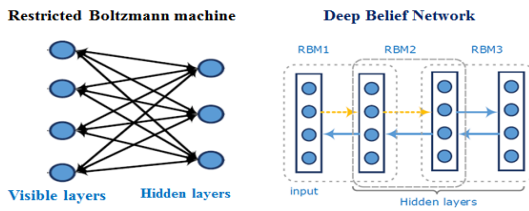


Fig. 2. Deep belief network architecture.

2.3. Convolutional Neural Networks (CNNs)

The convolutional neural network (CNN) is a multilayer, feed-forward neural network that uses perceptrons for supervised learning and to analyze data. It is used mainly with visual data, such as image classification. CNN architecture is different from other neural networks. The hidden layers in CNN contain convolutional layer, subsampling layer (pooling layer) and a fully connected layer (Fig. 3). Characteristically, CNN start with convolutional layer that accepts data from input layer. The convolutional layer is responsible for convolution operations having few filter maps of same size. The convolutional layer uses the convolution operation to achieve the weight sharing while the subsampling is used to reduce the dimension [6, 7].

Following the convolutional layer, a subsampling (or pooling) layer is usually used to reduce the dimension of the feature map. It can typically be realized by an average pooling operation or a max pooling operation. After the second stage, CNN uses a fully connected layer and then a softmax layer with output classes for classification and recognition.

During recent years, CNN has achieved great success in many applications such as image analysis, speech recognition, text understanding and so on [7].

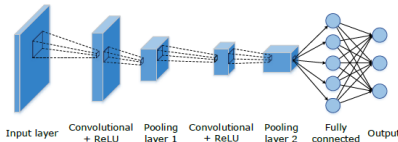


Fig. 3 Architecture of a CNN.

2.4. Recurrent Neural Networks (RNNs)

Recurrent neural network considered as another class of deep networks for unsupervised / supervised learning that is very powerful for modeling sequence data (e.g., speech or text). RNN learns features for the series data by a memory of previous inputs that are stored in the internal state of the neural network. The connections between neurons is constructed with a directed cycle (Fig. 4.).

Unlike traditional networks, where inputs and outputs are independent of each other, the recurrent neural network captures the dependency between the current sample with the previous one by integrating the previous hidden representation into the forward pass. From a theoretical point of view, the recurrent neural network can capture arbitrary-length dependencies. However, it is difficult for the recurrent neural network to capture a long-term dependency because of the gradient vanishing with the back-propagation strategy for training the parameters. To tackle this problem, some models, such as Long Short-Term Memory, have been presented by preventing the gradient vanishing or gradient exploding [7]. The recurrent neural network and its variants have achieved super performance in many applications such as natural language processing, speech recognition and machine translation.

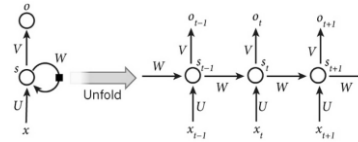


Fig. 4 Architecture of a RNN

3. Deep learning in Big Data analytics

The big data application process generally includes stages such as data generation, data management, data analytics, and data application. Big data analytics, which is considered the most important phase in the whole chain, refers to the process of discovering patterns from data. In this stage, there are several challenges (such as high dimensionality, scalability of algorithms, fast moving streaming data, noisy and poor quality data and so on), which is made big data analytics much more difficult and complicated than normal-sized data analytics [10].

In this section have been provided the analysis and discussions about the challenges and possible solutions of deep learning for big data analytics.

3.1. Complex data representation

Big Data is usually collected from different domains which consists of multiple modalities. Each modality has a different representation, distribution, scale, and density. For example, text is usually represented a discrete word-count vectors, but an image is represented by real values of pixel intensities [11]. The using of existing methodologies for the processing of such data is almost impossible. The solution of this problem is possible owing to the integration of heterogeneous data.

Deep Learning is more fitting for heterogeneous data integration due to its potentiality of learning variation factors of data and providing abstract representations for it. Deep learning has been demonstrated to be very effective in integrating data from different sources [3]. Some multi-model deep learning models have been proposed for heterogeneous data integration.

For example, Ngiam et al. [12] developed a multi-modal deep learning model to learn representations by integrating audio and video data. Srivastava and Salakhutdinov [13] developed a multimodal Deep Boltzmann Machine (DBM), for text data and image objects feature learning.

Ouyang et al. [14] presented multi-modal deep learning model, called multi-source deep learning model aims to learn non-linear representation from different information sources. In this model each source of information is used as input data for the two hidden layers deep learning model. Extracting features separately are then combined for joint representation.

Generally, though the architecture of the proposed multi-modal deep learning models is different, their ideas are similar. In particular, multi-modal deep learning models firstly learn features for single modality. Then learned features are combined as the joint representation for each multi-modal object. These models have been achieved more superior productivity than traditional deep neural networks for heterogeneous data feature learning. However, these models combine the learned features of each modality in a linear way. So they are far away effective to capture the complex correlations over different modalities for heterogeneous data. In order to eliminate this problem, Zhang et al. [15] presented a tensor deep learning model, called deep computation model, for heterogeneous data.

3.2. Super-high dimensionality

Big data in specific domains is often super-high dimensional. Generally, with the increase of the data dimension, the required amounts of time or memory go up exponentially. The problem is that existing machine learning and data mining algorithms are not well scalable to high-dimensional data (such as, images), or are not computationally efficient.

Chen et al. [16] developed marginalized stacked denoising autoencoders (or mSDAs) which scale effectively for high-dimensional data and is computationally faster than regular stacked denoising autoencoders (SDAs). This approach marginalizes noise in SDA training and therefore does not require other optimization algorithms to learn parameters.

Zhang et al. [17] proposed a new tensor-based representation algorithm for image classification. The algorithm is realized by learning the parameter tensor for image tensors which the algorithm preserved the spatial information of image.

Convolutional neural networks also can scale up effectively to high-dimensional data. On ImageNet dataset with 256×256 RGB images, CNNs produced state-of-the-art results [4]. For instance, Krizhevsky et al. [18] trained one of the largest Deep Convolutional Neural Networks (DCNN) to classify ImageNet LSVRC-2010 contest which comprises 1.2 million high-resolution images belonging to 1000 different image classes. It is one of the most well-known CNN architectures for classification. This large DCNN consists of 650,000 neurons with 60 million parameters and eight layers.

Maggiori et al. [19], proposed an end-to-end framework for the dense, pixel-wise classification of satellite imagery with convolutional neural networks.

The above Deep Learning algorithms for Big Data Analytics involving high dimensional data are not sufficient, and requires new methods for better performance of DL techniques to handle high-dimensional data.

3.3. Unscalable computation ability

A big dataset often includes a large number of attributes and many class types of samples, so some frequently used data mining and machine learning algorithms, is not work well. In order to learn features and representations for large amounts of data, some large-scale deep learning models have been developed. They can be grouped into three categories, such as *parallel deep learning models*, *GPU-based implementation*, and *optimized deep learning models* [7].

Existing deep learning systems commonly use data or model parallelism, but unfortunately, these strategies often result in suboptimal parallelization performance. Z. Jia et al. [20] proposed FlexFlow, a deep learning system that automatically finds efficient parallelization strategies for DNN applications. Authors evaluate FlexFlow with six real-world DNN benchmarks on two GPU clusters and show FlexFlow significantly outperforms state-of-the-art parallelization approaches.

Dean et al. [21] determined the possibility of training a deep network with billions of parameters using tens of thousands of CPU cores. Authors have developed a software framework called DistBelief that can utilize computing clusters with thousands of machines to train large models. DistBelief needs 16 thousand CPU cores to train a large deep learning model with 10 million images and billion parameters.

Sun et al. [22] presented techniques to accelerate distributed training of DNN on GPU clusters. They used two clusters: a cluster with 16 machines, each having 8 Pascal GPUs and a cluster with 64 machines, each having 8 Volta GPUs.

Coates et al. [23] deployed a less expensive cluster of (GPU) servers and also Commodity OFF-The-Shelf (COTS) HPC technology with a high-speed communication network to coordinate distributed computations. This system is capable to training for 1 billion parameters networks on just 3 machines in a few days and is capable scaling up to 11 billion parameters with 16 machines. Therefore, this system is affordable for everyone who wishes to explore large scale systems.

Novikov et al. [24] proposed a tensorizing learning model based on the tensor-train network. Authors converted the neural network to the tensor format to use the tensor-train network to compress the parameters. This method could reduce the computational complexity and improve the training efficiency in the back-propagation procedure.

There is a need to develop new algorithms for scalable deep learning which make it suitable for high dimensional data processing and analysis.

3.4. Fast moving streaming data

One of the challenging aspects in Big Data Analytics is dealing with streaming and fast-moving input data. The data stream is generated at an extremely fast speed, and its distribution characteristics are in high-speed dynamic changes, which must be processed in real time. Deep learning to handle streaming data, as there is a need for algorithms that can deal with large amounts of continuous input data. In recent years, a lot of incremental learning methods have been presented for high-velocity data feature learning.

Zhou et al. [25] proposed an incremental feature learning algorithm to determine the optimal model complexity for large-scale datasets based on the denoising autoencoder. The model quickly converges to the optimal number of features in a large-scale online setting. In addition, the algorithm is effective in recognizing new patterns when the data distribution changes over time in the massive online data stream. Calandra et al. [26] demonstrated Adaptive Deep Belief Network to learn from online, nonstationary stream data.

Y. Li and et al. [27] proposed an incremental high-order deep learning model based on parameter updating and structure updating to meet the requirements of dynamic big data online analysis and real-time processing. The model has the ability to incrementally learn the characteristics of new data online, also retains the ability to learn the original data features, and real-time processing of dynamic data streams.

3.5. Noisy and poor-quality data

There are a huge number of noisy objects, incomplete objects, inaccurate objects and imprecise objects in Big data. This low-quality data is widespread in Big data. For example, there are over 90% missing attribute values for a doctor diagnosis in clinic and health fields. Some traditional learning algorithms have obviously not been valid for processing the data with 90% missing values. In the past few years, some methods have been proposed to learn features for poor-quality data.

Wang and Tao presented a non-local auto-encoder model to learn reliable features for corrupted data [280]. The model achieved high performance in image denoising and restoration. Mao et al. [29] proposed a very deep fully convolutional auto-encoder network for image restoration. Since this method is based on convolutional operations, its main limitation is the local nature of the extracted features.

In [30], a deep convolutional neural network has been proposed for image denoising, where residual learning is adopted to separating noise from noisy observation.

Recent methods based on CNNs can only operate local similarities and they are incapable to capture non-local similar to itself patterns, which have been highly successful in model-based methods. In order to exploit both local and non-local similarities, in [31], has been proposed a graph-convolutional neural network, to perform image denoising. This method provides the best visual quality, recovering finer details and producing fewer artifacts.

4. Results and discussion

Researches show that significant progress has been obtained in the application field of deep learning algorithms in Big data analytics. DL sufficiently simplifies solution of Big data analytics problems as analysis of large data volumes, semantic indexing, data tagging, information retrieval, classification and prediction. At the same time, deep learning has achieved limited progress in the field of stream data and low-quality data processing, model scaling, distributed computing, and high-scale data processing. Below have been outline several open issues and research trends.

1) Continuous increasing of volume of big data makes it necessary to create more large-scale deep learning models. Such large-scale deep learning models that can be trained for Big Data

may no longer be effectively trained, depending on the available techniques and computing power. It is important to create new learning structures and computing infrastructures in the future to solve this problem.

2) Modern multi-modal deep learning models simply combine in a linear form the learned features of each modality. This often does not lead to the necessary results. There is need to investigate the effective fusion ways of learned features to improve the productivity of multi-modal deep learning models. At the same time, deep computational models have a large number of parameters that caused their high computational complexity. There is a need to researches in the field of reducing the computational complexity of deep computational models.

3) Most of the integrated learning algorithms that based on updates of parameters or structure are effective only for a hidden, layer, traditional learning models. There is a need to research of the application possibilities of integrated learning algorithms to deep learning models and deep architectures.

4) It is important to investigate reliable deep learning models for low-quality data in the near future, due to the rapid growth of low-quality data.

5) There is a need to develop new parallel and distributed algorithms/frameworks for scalable deep learning models.

5. Conclusion

In this paper has been investigated how deep learning algorithms and architectures are used to solve Big Data analytics problems. An overview of significant literature according to the application of Deep Learning in different domains showed that Deep Learning has the potential opportunities to the solving of many analytics and learning challenges faced by Big Data analytics unlike traditional machine learning methods. But while Big Data offers enough training objects for deep learning, it creates problems for large scale, heterogeneity, noisy labels, and non-stationary distribution, among many others. In order to realize the full potential of Big Data, we need to address these technical challenges with new ways of thinking and transformative solutions. For this reason, there is need for extensive investigates in the field of deep learning the future.

References

1. Alguliyev, R. "Big Data" phenomenon: Challenges and Opportunities. - Problems of Information Technology, vol. 10, no. 2, 2014, pp. 3-16. (Alguliyev R., M. Hajirahimova)
2. Alguliyev, R. Current scientific and theoretical problems of big data. - Problems of Information Society, vol. 10, no. 2, 2016, pp. 34-45. (Alguliyev R., M. Hajirahimova. A. Aliyeva)
3. Chen, W. Big Data Deep Learning: Challenges and Perspectives. - Access IEEE, vol. 2, 2014, pp. 514-525, (Chen W., X. Lin)
4. Najafabadi, M. Deep Learning applications and challenges in Big Data analytics. - Journal of Big Data, vol.2, no.1, 2015, pp.2-21. (Najafabadi M., F. Villanustre, T. Khoshgoftaar, N. Seliya.)
5. N. M. Elaraby, M. Elmoghy, and Sh. Barakat, "Deep Learning: Effective Tool for Big Data Analytics." International Journal of Computer Science Engineering, vol.5, no.5, pp. 254-262, 2016.
6. Jan, B. Deep learning in big data Analytics: A comparative studym. - Computers and Electrical Engineering, vol.7, no. 24, 2017, pp. 1-13. (Jan B., H. Farman, M. Khan, M. Imran)
7. Zhang, Q. A survey on deep learning for big data. - Information Fusion, vol. 42, 2018. pp. 146-157. (Zhang Q., L. Yang, Z. Chen)
8. Liu G. A Stacked Autoencoder-Based Deep Neural Network for Achieving Gearbox Fault Diagnosis. - Mathematical Problems in Engineering, 2018, no. 5, pp. 1-10. 2018. (Liu G., H. Bao, B. Han)
9. Hinton, G. A fast learning algorithm for deep belief nets. - Neural computation, vol. 18, no. 7, pp. 1527-1554, 2006. (Hinton G., S. Osindero, Y.-W. Teh)
10. Wang, X. Learning from Uncertainty for Big Data: Future Analytical Challenges and Strategies. - IEEE Systems, Man., & Cybernetics Magazine, April 2016, pp.26-32 (Wang, X., Y. He)
11. Zheng, Y. Urban Computing, Cambridge, The MIT Press, 2018, 609 p. (Zheng, Y.)
12. Ngiam, J. Multimodal deep learning. - 28th Inter. Conference on Machine Learning. ACM, June 28 - July 2, 2011, pp. 689-696.
13. Srivastava, N. Multimodal learning with deep boltzmann machines. - Advances in Neural Information Processing Systems, MIT, 2012, vol. 25, pp. 2231-2239.
14. Ouyang, W. Multi-source deep learning for human pose estimation. - IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2014, pp. 2337-2344.
15. Zhang, Q. Deep computation model for unsupervised feature learning on big data. - IEEE Transactions on Services Computing, vol. 9, 2016, pp. 161-171. (Zhang Q., L. Yang, Z. Chen.)
16. Chen, M. Marginalized denoising autoencoders for domain adaptation. - 29th International Conference in Machine Learning, Edinburgh, 2012, Scotland
17. Zhang, J. Semi-supervised tensor learning for image classification. - Multimedia Syst., vol. 23, no. 1, 2017, pp. 63-73. (Zhang J., Y. Han, J. Jiang)
18. Krizhevsky, A. Imagenet classification with deep convolutional neural networks. - Advances in Neural Information Processing Systems. Curran Associates, Inc. 2012, vol. 25. pp 1106-1114.
19. Maggiori, Y. Convolutional neural networks for large-scale remote-sensing image classification. - IEEE Trans. Geosci. Remote Sens., vol.55, no.2, 2017, pp. 645-657. (Maggiori Y., G. Tarabalka, P. Charpiat)
20. Jia, Z. Beyond data and model parallelism for deep neural networks. - arXiv:1807.05358v1 [cs.DC] 14 Jul 2018, pp.1-15. (Jia, Z., M. Zaharia, A. Aiken)
21. Dean, J. Large scale distributed deep networks. - Proceedings of NIPS, 2012, pp. 1232-1240.
22. Sun, P. Optimizing Network Performance for Distributed DNN Training on GPU Clusters: ImageNet/AlexNet Training in 1.5 Minutes. - arXiv preprint arXiv:1902.06855, 2019. (Sun P., W. Feng, R. Han, S. Yan, Y. Wen)
23. Coats, A. Deep learning with COTS HPC systems. - J. Mach. Learn. Res., vol.28, 2013, pp. 1337-1345. (A. Coats A., B. Huval, T. Wng, D. Wu, A. Wu)
24. Novikov, A. Tensorizing neural networks. - Advances in Neural Information Processing Systems, MIT, 2015, pp. 442-450.
25. Zhou, G. Online incremental feature learning with denoising autoencoders. - International Conference on Artificial Intelligence and Statistics. JMLR.org. 2012, pp 1453-1461.
26. Calandra, R. Learning deep belief networks from non-stationary streams. - Artificial Neural Networks and Machine Learning-ICANN 2012. Springer, Berlin Heidelberg. 2012, pp. 379-386.
27. Li, Y. Online Real-Time Analysis of Data Streams Based on an Incremental High-Order Deep Learning Model. - IEEE Access, vol. 6, 2018, pp. 77615 - 77623, (Li Y., M. Zhang, W. Wang)
28. Wang, R. Non-local auto-encoder with collaborative stabilization for image restoration. - IEEE Transactions on Image Processing, vol. 25, no. 5, 2016, pp. 2117-2129. (Wang, R., D. Tao)
29. Mao, X. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections - Advances in Neural Information Processing Systems, vol. 29, 2016, pp. 2802-2810.
30. Zhang, K. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. - IEEE Transactions on Image Processing, vol. 26, no. 7, 2017, pp. 3142 - 31555, 2017. (Zhang K., W. Zuo, Y. Chen, D. Meng, L. Zhang)
31. Valsesia, D. Image denoising with graph-Convolutional Neural Networks. - 2019 IEEE International Conference on Image Processing (ICIP), 22-25 Sept. 2019, pp. 2399 - 2403.

Concept of parametric design and fast 3D printing of individual shoes

Valentina Vladimirovna Kostyleva¹, Olga Borisovna Konovalova²

Art Modeling Design and Technology of Leather Products department, Russian State University named after A.N. Kosygin, Moscow^{1,2}
kostyleva.vv@mail.ru, o.b.konovalova@gmail.com

Abstract. A personalized approach to printing individual shoes is possible due to the availability of laser scanning of the feet and professional interpretation of the results. Based on preliminary diagnostics of the condition of the musculoskeletal system in areas of high pressure in the feet, it is possible to identify problems associated with pain in the joints of the lower extremities and lumbosacral, abnormalities and asymmetries of their structure. It is also possible to timely relieve ankle, knee, femoral joints, lumbosacral spine, correct posture and gait, reduce joint pain, prevent surgery on joints of the lower extremities and lumbosacral spine, monitor the dynamics of changes in the rehabilitation period.

The task of effective prevention and early diagnosis of problems of the musculoskeletal system can be solved with the targeted popularization of individual shoes and orthopedic insoles among the population, as well as its accessibility thanks to additive technologies, databases of parametric digital models, an established production and technological cycle and a flexible client-oriented service, including protocols for remote and virtual interaction between production and the end user.

However, shoes printed on a 3-D printer, so far, are not attractive to consumers or manufacturers. Mass production of shoes is still much faster and cheaper, although initially it was assumed that the introduction of additive technologies will significantly reduce the cost and time costs, but so far this has not happened for several reasons. In our country, there is a shortage of specialists in digital technologies, 3D-modeling and printing. In addition, as elsewhere in the world, equipment for fast 3D printing has a very high cost and only giants such as Adidas or Nike can create real production of printed products. However, it is a matter of time. The area of additive technologies is attracting more and more medium and small firms, individual designers, because 3D printing allows you to achieve unexpected and revolutionary results, as well as provide customized solutions at an affordable price.

KEYWORDS: DIGITAL TECHNOLOGIES, PARAMETRIC DESIGN, 3D-PRINTING, INDIVIDUAL SHOES.

1. Introduction

1.1. An increase in the number of people with pathologies of the musculoskeletal system.

A more in-depth and personalized approach to printing individual shoes is possible due to the availability of laser scanning of the feet and professional interpretation of the results. Based on preliminary diagnostics of the condition of the musculoskeletal system in areas of high pressure in the feet, it is possible to identify problems associated with pain in the joints of the lower extremities and lumbosacral, abnormalities and asymmetries of their structure. It is also possible to timely unload the joints of the lower extremities and the lumbosacral spine, reduce pain in these areas, prevent operations on the joints of the lower extremities and the lumbosacral region of the spine, identify the dynamics of changes in the rehabilitation period, and adjust posture and gait [1].

1.2. The high cost of printed shoes due to monopolies in this market segment.

Shoes printed or connected on a 3D printer in our country are gaining popularity rather slowly and in a limited segment - for sports (Nike, Adidas, New balance, etc.). Initially, it was assumed that additive technologies would make the cost of such shoes low and affordability high. However, this has not yet happened (the cost of a pair of 3D sneakers of the above brands can reach from 300 - 600 dollars in retail). As a result, a third problem arises.

1.3. The growing inaccessibility of high-quality individual and orthopedic shoes.

According to Rosstat's research on the living conditions of Russians dated April 5, 2019, more than 35% of Russian families cannot afford to buy one pair of shoes for each season. It is only about everyday shoes, excluding categories of the population with orthopedic problems.

2. Background and means to solve the problem

2.1. The situation in the segment.

Despite the fact that such giants of the shoe industry as Adidas, Nike, Under armor, New balance, Reebok have long and successfully used technology for fast and economical 3D printing to produce individual parts of the upper and lower shoes. For example, layer-by-layer photopolymerization of resins in a bath (or stereolithography) - SLA, DLS, etc. However, mass production of shoes is still much faster and cheaper.

This situation occurs for several reasons:

- In our country, there is a shortage of specialists in digital technologies, 3D-modeling and printing.
- We, as well as around the world, have equipment for fast 3D printing at a very high cost, and only giant companies such as, for example, the American concern Carbon can create real mass production of printed products.

However, it is a matter of time. The area of additive technologies is attracting more and more medium and small companies, individual designers, because 3D printing allows you to achieve unexpected and revolutionary results, as well as provide customized solutions at an affordable price.

In addition, engineering thinking is evolving, leading to the introduction of new design methods, such as parametric design.

2.2. Fashion tech.

The beginning of the XXI century is characterized by a large number of scientific studies and discoveries. The technological breakthrough did not pass by the world of fashion. The development of the Fashion Tech trend in the variety of its manifestations is associated with the enormous opportunities that modern digital technologies provide the industry with. The origins of these phenomena are the 20th century, which has overturned the generally accepted notions of life on our planet, its forms and development. Such theories, for example, as chaos theory, complexity theory, quantum field theory, etc., have formed a new synergistic picture of the world. Researchers have identified such qualitative changes in the paradigm of public consciousness as the transition: from a picture of the sustainable development of society to a vision of its metastability; from order to chaos; from evolution to co-evolution of complex systems; from independence to dependence; from dimensionality to proportionality or fractal replication by the world of itself. Parametric design, which came into practical use at the beginning of the XXI century, basically reflects the entire spectrum of emerging ideas about the world as a complex system of interdependent elements, changes in one of the parameters of which leads to a change in the whole system, as well as the basic algorithms that lie in basis of this system.

2.3. Parametric design.

In the process of designing a new product, or customizing an existing one, many tasks arise for combining data of different generation and characteristics that are dependent and independent from each other. Thus, the process of designing a complex product can be delayed, and making changes to an existing project can give rise to a series of self-replicating errors, significantly complicating

the process of further work. Especially, in a situation with a predominance of "manual" operations in the project.

The undoubted advantage of automation of technological processes is a significant reduction in time spent on product development. However, specialized CAD systems require the involvement of experienced technical professionals, who, unfortunately, are not enough to meet the needs of the industry. Thus, specialized CAD systems are ineffective or inaccessible to small enterprises or private practitioners. At the same time, the active expansion of additive technologies in all branches of industry allows us to implement the most ambitious ideas in the material, which, of course, attracts the attention of specialists from various fields (designers, marketers, etc.).

Against this background, parametric design seems to be attractive, allowing you to create algorithms for objects or processes of any complexity based on the analysis of data arrays, as well as their interaction in real time, which serves as a universal mechanism for creating both individual objects and entire systems. The scalability of projects is the most important property of this approach, which enables the specialist to manage the complexity of the project, to build any hierarchical structure [2, 4-7].

One of the most important properties of the parametric approach is its adaptability to environmental factors. Using the method allows you to create objects consisting of groups of autonomous elements. Some of these groups respond to the influence of external factors by changing the properties of the groups associated with them and, as a consequence, the system as a whole. For example, in response to a change in pressure force, the texture density and surface pattern of the material change.

2.4. "Rhinoceros & Grasshopper."

Universal parametric design tools such as Rhinoceros and Grasshopper allow a creative specialist to realize ideas that are difficult to understand, easily translating them into an algorithm language and expressing them in a visual modeling environment. It is the creator of the concept that can choose from a variety of design parameters that are most important for the full implementation of his plan. Also, any number of changes are made quickly and clearly, without requiring the replication of repeated operating cycles.

For example, the motive of the collection is a cellular fractal structure, in which each element, having its own set of parameters, is part of another cellular fractal structure, being one of the parameters for its element. Generally speaking, this is a kind of grid located inside another grid that is inside the third grid, etc. And, thanks to additive technologies, such as 3D printing on fabric, collections of this kind, including scalable parametric motifs, are now not uncommon (Atlantis Plato, Alexander McQueen).

The synergy of the designer's creative intent and capabilities of digital design technologies allows us to take a fresh look at the growing role of form and its embodiment in the modern fashion industry. The development of the natural sciences shows us the evolution of complex objects of animate and inanimate nature, as, kind of, derived from the processes of fusion, division and transformation, fueling the interest of scientists and designers in biomorphic and topomorphic objects. Complex objects of living nature undergo significant transformations under the influence of many laws throughout the life cycle, providing a vivid example of the parametric principles of shaping in the environment.

2.5. "Wearable technologies".

Such project logic implies reaching a fundamentally different level of creating clothes and shoes - the level of formation of the vital membranes of the human body, its "small infrastructure". According to some scientists, in the near future it will be possible to design complex objects and systems in those areas of life where shells created or built according to the principles of molecular synthesis will arise directly at the object of application of their functions, and shape and design correction will occur by changing the software code. Outdated or faulty elements of such an infrastructure will be able to be updated or self-repaired according to the same algorithmic principles. Futurologists believe that clothing and the environment, united by a unity of functions, will

have their own ecosystem, harmoniously integrated into the surrounding nature.

"Wearable technologies" - clothing (from underwear to shoes and accessories), is directly related to such a thing as metamaterials, which can change their properties by changing the spatial structure. Thanks to them, dimensionless textiles became a reality. So, the Royal College of Arts recently introduced the development of children's dimensionless clothing. Increasingly, they are announcing products based on nanotissues that respond to light, ambient temperature, human condition, and the nature of their activities. Spray clothing technology has already been developed by Imperial College London. Due to the availability and versatility of properties, the material is predicted a great future in various fields of application, from fashion to medicine.

Modern materials containing natural fibers, such as silk, as well as nanoparticles of an electrical conductor, flexible sensors, etc., make smart fabrics a reality. Sports and physical activity are not always equivalent to "health" and "safety". Therefore, intelligent sportswear for diagnosis and prevention is included in everyday life. However, the inclusion of technological components in clothing requires a change in the thinking of both manufacturers and consumers. Therefore, for integration into our everyday life, "smart clothes" is becoming more and more like everyday clothes.

The increase in the density and concentration of the urban environment, the acceleration of social processes, the need to provide protection and comfort at various standards of consumption, forces engineers and designers to use new adaptive parametric technologies and materials.

3. The solution to the problem

We see that the problems of increasing the availability of individual shoes, as well as effective prevention and early diagnosis of disorders of the musculoskeletal system, can be solved by:

- targeted popularization among the population of individual printed shoes and orthopedic insoles,
- the advent of inexpensive domestic equipment for 3D printing,
- established production and technological cycle;
- flexible client-oriented service, including protocols for remote and virtual interaction between production and the end user.

3.1. Hypothesis - the production of affordable individual shoes printed on a 3D printer is possible due to the availability of available equipment, supplies and software, as well as:

- a well-developed base of parametric models;
- algorithmic design in real time for customization of models from this base;
- mobile client service and VR / AR technologies (mobile offices, virtual fitting and mobile application).

3.2. Goal

The purpose of the work is to increase the efficiency of technological processes in the domestic shoe industry and the quality of manufactured products.

3.3. Tasks of work:

- create an algorithm for converting graphic data of foot pressure force to the surface obtained as a result of scanning into a digital model of upper and lower parts of shoes made of 3D-printed material with a reconfigurable relief structure, the density and height of which depend on the pressure force;
- analyze the properties and behavior of the digital model of the upper and lower parts of the shoe;
- create prototypes and samples of a digital model;
- analyze the change in their elastic properties depending on the materials used;
- implement the results in the form of a working module of the technological complex of parametric design and quick 3D printing of individual shoes in the format of a mobile salon.

4. What is planned to be achieved:

1. Create an effective algorithm for converting the data of scanning the pressure force of the foot onto the supporting surface, into a digital model of the upper and lower parts of the shoe with a reconfigurable relief structure for quick printing of individual shoes in the format of mobile salons;

2. Create a database of digital models of upper and lower shoe parts;

3. Create a work module for the technological complex of parametric design and fast 3d printing of individual shoes in the format of a mobile salon.

5. Methods and tools for implementation:

Methods of data collection and analysis:

- This work analyzes sources on the topic, the results of laser foot scans, conducts experimental studies to create digital models of materials with a reconfigurable structure and simulate their behavior, as well as studies of real materials (sheets and filaments).

- Prototypes are created, and their elastic behavior is analyzed taking into account the properties of real materials and boundary conditions.

Tools and equipment:

- BY,
- 3d printers,
- results of 3D scans,
- sheet materials,
- filaments.

6. Proposed solutions.

6.1. The optimal network structure of the 3D shoe printing mobile showrooms, in our opinion, should include:

1. One head office, in which there is one designer, one technologist, one orthopedist, as well as a base of models, blocks, soles, materials.

2. N-number of mobile salons equipped with all necessary equipment, in which there is one operator who controls a 3D scanner, a 3D printer, communicates with customers, courier service and head office, receives and issues orders.

3. Logistics department-warehouse, connecting the network with the supply center, delivering orders to customers involved in transportation.

4. The client communicates with the designer and orthopedist remotely. Mobility allows you not to get attached to the location and move the salon to the points of maximum concentration of the target audience (various social centers, events, etc.)

6.2. The basic technological cycle will consist of the following steps:

1. The only appeal: the client comes to the mobile salon, where he selects a model from the catalog or develops an individual one together with a specialist who conducts a topological analysis of the distribution of foot pressure on the supporting surface and interprets the results of the study. The client receives a digital scan model of their feet on digital media; if necessary, additionally consults with the designer and orthopedist.

2. After receiving the order, the specialist conducts a more detailed

morpho-functional analysis of scan results.

3. Next, the selection of materials for printing is carried out depending on the configuration of the three-dimensional structure and the number of its layers, the pressure value.

4. The algorithm of the interdependence between the pressure zones and the density of the relief pattern is established.

5. Relief modeling is carried out, the vector model is converted to a solid model to be sent to the preparation program for 3d printing.

6. A new design for the details of future shoes is being created. If necessary, customization of the finished design solution from the catalog is performed.

7. Directly in the salon or in the central office (depending on the complexity of the project), 3D printing of the upper and lower parts of the shoes is performed, their connection.

All clarifications with the client are carried out remotely through instant messengers or the company's information portal in 24/7 mode. It is also possible to develop a virtual application for trying on a product before delivery or during production / alteration.

Shoes are delivered to the client by courier or mail, (pickup possible).

6.3. Expansion Technology Cycle:

1. Products can be equipped with additional tools, for example, nanospray-preventive disposable socks, for gentle and comfortable use of shoes and protect the foot from scuffs and calluses.

2. Depending on the polymers used, micro-offices may provide services for collecting used 3D shoes for further processing.

3. As a marketing move, as well as a scaling element, you can offer customers a kind of "designer" of individual elements (soles, upper parts, individual pads, 3D models of feet), 3D handles and the corresponding polymers for independent amateur customization of individual shoes at home.

4. In addition, it is possible to integrate systems of elastic sensors into shoes for timely diagnosis and prevention of diseases according to the position of the body in space and the difference in pressure of the foot sections to the surface. And with the help of its own mobile application, "smart shoes" will instantly inform its owner about an error in the distribution of load on the joints during training or in ordinary life.

7. The expected result and directions of its practical use.

It is planned to create an effective algorithm for converting graphic data of the pressure force of the foot to the surface, obtained as a result of scanning, into a digital model of the upper and lower parts of shoes with a reconfigurable relief structure, the density and height of which depend on the pressure force, in order to quickly print individual shoes in the format of mobile salons .

The final product, in essence, is a complex consisting of prompt customer service, affordable technology, environmentally friendly materials and high-quality individual shoes for all categories of the population at an affordable price.

8. A wide target audience.

Accessibility, efficiency, the possibility of cheapening by ordering standard models from the catalog, unlimited possibilities of customization, as well as providing quality service to customers with special needs and orthopedic problems make the product suitable for all categories of the population.

9. A scalable commercialization plan.

- The modular structure with reduced labor costs enables both remote service and business scaling to the regions, regardless of their economic and environmental conditions, as well as the sale of franchises.

- In addition, the development of hardware, software, mobile applications, integrable "smart" blocks, materials, both for business and for the end user, provides an opportunity to develop in these areas.

Separately, it is necessary to mention the target audience with special needs.

The estimated competitive advantages of the methodology for the target audience with problems of the musculoskeletal system:

- unloading of the ankle, knee and femoral joints, lumbosacral spine, and, as a result, reduction of pain in them;
- posture and gait correction;
- the ability to track the dynamics of changes;
- prevention of surgical interventions on the joints of the lower extremities and lumbosacral region [3].

10. Explanatory Note

11.1. Purpose and scope:

- for a wide consumer of affordable individual shoes;
- for the target audience with problems of the musculoskeletal system.

11.2. Specifications.

Algorithm for translating graphic data obtained by scanning the foot into digital models of the upper and lower parts of shoes with a reconfigurable relief structure, the density and number of layers of which depend on the foot pressure on the surface.

11.3. TEP.

Based on assessments of the results of the practical implementation of parametric design tools in related fields, we can assume a significant reduction in the following indicators:

- 12-15% reduction in the cost of paying for production downtime, as well as energy consumption by reducing downtime in the technological cycle, its acceleration and simplification;
- the residual carbon footprint from shoe production will be reduced by 30-35%;
- the development and customization of the product will be reduced by 50-60% due to the base of parametric digital models.

11. Conclusion

Currently, the work is at the stage of developing an algorithm for converting graphic data obtained as a result of scanning into digital models of upper and lower parts of shoes with a reconfigurable relief structure, the density and height of which depend on the force of the foot pressure on the surface. But it is already possible to draw a preliminary conclusion that the

unification of this process, with the receipt of a library of such samples, will several times reduce the time spent on the development of prototypes and their customization.

Directions for further research:

- Parametric programming algorithms;
- Latest software and algorithmic design technologies;
- 3D printing technology;
- 3D scanning technology;
- Consumables, filaments.

12. Literature

1. Li and Tanaka. Feasibility study applying a parametric model as the design generator for 3D-printed orthosis for fracture immobilization. *3D Printing in Medicine* (2018) №4.
2. Steven S. Skiena. *The Algorithm Design Manual*. Second Edition. Springer. 2008.
3. Bekk N. V., Kostyleva V. V., Lapina T. S. Customizing of the techniques used for designing of the orthopedic footwear. *New Technologies, Development and Application II*. Springer Nature Switzerland AG 2020 I. Karabegović (Ed.): NT 2019, LNNS 76, pp. 803–807.
4. Miura and Tachi. "Synthesis of Rigid-Foldable Cylindrical Polyhedral." (*Symmetry: Art and Science*, International Society for the Interdisciplinary Study of Symmetry, Gmuend),
5. Yasuda and Yang. "Reentrant Origami-Based Metamaterials with Negative Poisson's Ratio and Bistability." (*Physical review letters* 114:18 (2015), 185502),
6. Yasuda, Yein, Tachi, Miura, and Taya. "Folding behaviour of Tachi-Miura polyhedron bellows." (In *Proc. R. Soc. A*, 469, 469, p. 20130351. The Royal Society, 2013),
7. Yasuda, Chong, Charalampidis, Kevrekidis, and Yang "Formation of rarefaction waves in origami-based metamaterials." (*Phys. Rev. E* 93 (2016), 043004. doi:10.1103/PhysRevE.93.043004)

An experimental study on energy generation from photovoltaic-thermal hybrid systems

Danijela Kardaš^{1,*}, Bojan Knežević¹, Milan Pupčević¹

University of Banja Luka, Faculty of Mechanical Engineering, Bosnia and Herzegovina¹
danijela.kardas@mf.unibl.org

Abstract: Solar energy is one of the leading renewable energy sources because of its great potential, affordability of the individual system prices, zero emission and no noise emission. 21st century will be based on renewable energy sources, with strict environmental measures, primary energy use will mainly consist of a combination of (different) clean energy sources among which solar energy will play a leading role. Two technologies for harnessing the energy of the sun are photovoltaic cells (PV) and solar collectors. Photovoltaic cells (PV) exploit the visible part of the spectrum while solar collectors use the infrared part. Combining these two technologies into one provides a system that produces electricity and heat at the same time using all parts of the solar spectrum. These systems are known as photovoltaic thermal systems or PV/T. This paper presents the description of these systems and methods of application. Among this, this paper gives description of experimental PV/T installation at Faculty of Mechanical Engineering in Banja Luka.

KEYWORDS: PV PANELS, SOLAR COLLECTORS, PHOTOVOLTAIC-THERMAL SYSTEM, SOLAR ENERGY

1. Introduction

With population growth and the development of a global economy, the need for energy is steadily rising and fossil fuels are no longer as secure a source of supply as they were in the last century. From all of the above, renewable energy sources are the focus of both scientists and decision-makers with a view to ensuring the energy stability of states and protecting the environment. Solar energy is one of the leading renewable energy sources because of its great potential, affordability of the individual system prices, zero emission and no noise emission.

PV panels and solar collectors harness solar energy by covering a certain part of the spectrum of solar radiation. Photovoltaic cells (PV) exploit the visible part of the spectrum while solar collectors use the infrared part. Combining these two technologies into one provides a system that produces electricity and heat at the same time, utilizing all parts of the solar radiation spectrum. These systems are known as photovoltaic thermal systems or PV/T. PV/T systems can produce more energy per unit area compared to a single PV panel and solar collector [1]. This paper presents the description of these systems and methods of application. Among this, this paper gives description of experimental PV/T installation at Faculty of Mechanical Engineering in Banja Luka.

2. PV/T Systems: the working principle and classification

Only 15-20% of the solar radiation that falls on the PV module is converted to electricity while the rest heats the module [2]. In other words, more than 50% of the incident solar energy is converted as heat (after deducting the reflected portion). This may lead to extreme cell working temperature as much as 50°C above the ambient environment. The electrical efficiency of PV modules decreases as the operating temperature of the module rises. The electrical efficiency of the PV module depends on the material from which the module is made, as well as on the tilting angle of the panel [3], dust [4] [5], shadows [6] and the climate of a specific geographical area [7] [8] [9]. Significant impact on the reduction of PV module efficiency is the exposure to intense solar radiation and high temperatures for a long time [10].

The most critical periods for the electrical efficiency of PV modules are periods with high ambient air temperatures and low wind speeds [11]. Another problem that may occur due to the high temperature of the module is that photovoltaic cells can totally be burned [12]. From all of the above it is concluded that cooling the PV module is important and necessary. By cooling the solar cells with a fluid stream like air or water, the electricity yield can be improved. But conceptually the better design is to re-use the heat energy extracted by the coolant.

A hybrid system is a system that connects two or more energy conversion devices or one system to which multiple energy sources are integrated. The efficiency of a hybrid system is usually higher

than that of single technologies. In a hybrid PV/T system, the thermal energy of the panel is absorbed by the working fluid (coolant) by the heat conduction process or by convection, the panel temperature is reduced, thereby significantly increasing the electrical efficiency of the hybrid system. The advantages of PV/T systems are:

1. photovoltaic and solar collectors use different parts of the spectrum of solar radiation. Combining the two systems into one results in a more efficient use of the entire solar spectrum;
2. PV/T system requires less space for installation than separate installation of PV panels and solar collectors;
3. when implementing the system, the PV/T system requires less material than two separate systems;
4. electrical efficiency is increased by heat dissipation from the PV module;
5. the working fluid has a higher temperature due to the heat it absorbs from the photovoltaic cells;
6. Greater overall efficiency than separate systems.

PV/T systems are divided according to: type of collector, type of coolant, material of which PV is made, mode of flow of coolant. Fig. 1 shows the main features of a flat-plate PVT collector.

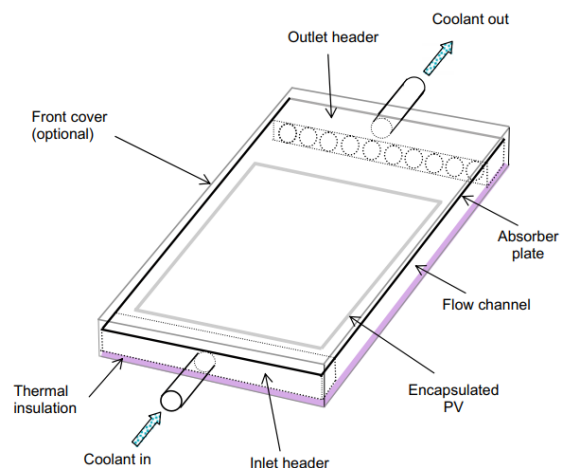
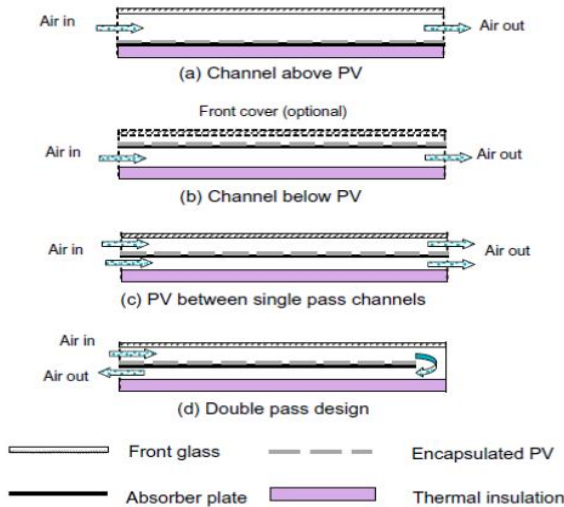
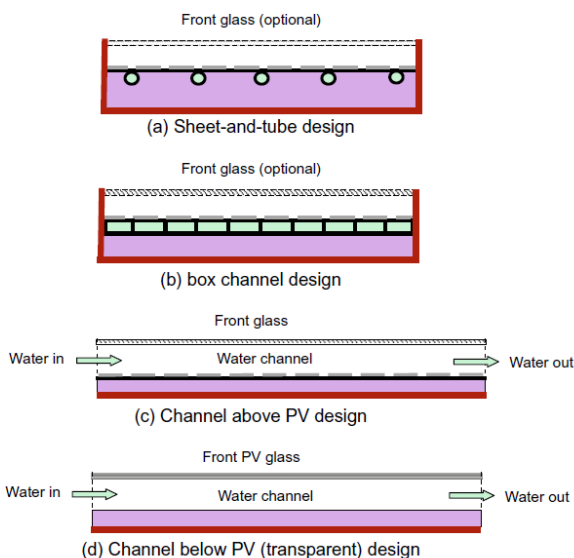
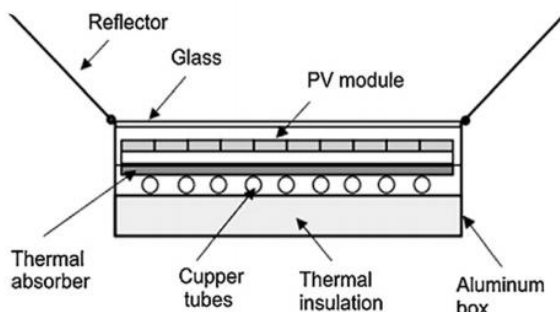


Fig.1. Main features of a flat-plate PV/T collector [13]

Fig. 2 shows the longitudinal cross-sections of some common air-type PV/T collector configurations whereas the cross-sectional views of some examples of the water-type designs are in Fig. 3.

Significant research has been carried out in recent years in the field of CPV/T systems (*Concentration photovoltaic/thermal system*) in order to achieve the highest efficiency of the system, ie higher electrical efficiency and higher output temperature from the system. However, when analyzing such systems, it must be borne in

mind that CPV/T systems are significantly more complex than flat plate PV/T systems, and that such systems are accompanied by a mirror-rotating device depending on the position of the sun. Photovoltaic cells in which the concentration of solar radiation is made are more expensive because of the working environment and exposure to higher radiation and high temperatures. Fig.4 is a simplified representation of the CPV/T collector.



The BIPV/T (Building Integrated PhotovoltaicThermal) system is the structural, architectural and aesthetic integration of the PV/T system into the building structure of the facility [15]. BIPV/T systems use the building envelope to collect solar energy to produce electricity and heat, providing an efficient way to reduce energy

consumption in the facility from other sources. Establishing a BIPV/T system as a standard in architecture but also as a functional component of a facility to replace conventional building materials for cladding would lead to a significant expansion of the use of the BIPV/T system. Given that 32% of the world's consumption is spent on buildings, this way of supplying the facility with heat and electricity is of interest for the development of NZEB buildings. The characteristics of the BIPV/T system are:

- the system is physically connected to the building;
- the system generates electricity;
- the system generates thermal energy that can be used directly in the building or through other technical systems.

Fig 5. Represents some implemented BIPV/T systems.



3. Experimental setup of PV/T system

Faculty of Mechanical Engineering in Banja Luka is implementing the research project "Thermodynamic Analysis and Mechatronic Synthesis of Solar Power Plants in Urban Areas" which is co-financed by the Ministry of Science and Technology Development, Higher Education and Information Society of Republika Srpska, Bosnia and Herzegovina. One of the project activities is experimental research of PV/T system in urban area. During research, it will be analyzed the possibility of obtaining electricity by conversion from solar energy as well as the thermal effects that accompany this phenomenon. Generated heat from PV panel will be used for water heating. Monitoring the trajectory of the Sun in order to obtain the maximum possible power in given climatic and meteorological conditions is a mechatronic type of a problem that is approached from the mechanical, energy/(thermal and electrical) and control aspects. The management mode is based on stored data that depends on geographical location, the use of light sensors or the methods of monitoring maximum output. The implementation of mechanical as well as control structures is different and is based on microcontrollers, programmable logic controllers, phase controllers and application of artificial intelligence. The idea of the project is to carry out scientific research that will answer the question of which solar monitoring system is the most energy efficient from the electrical and thermodynamic point of view, but also optimal in terms of the complexity of design, technical solution of mechanical part and price, and with innovative design acceptable for use in urban areas. During research it will be used monocrystal PV/T module and control monocrystal PV module (Solarfam Monocrystalline Solar Panels model SZ-100 - 36M). All measurement results (insolation, air temperature, precipitation, wind speed, data on different angles of positions of photovoltaic panels) will be used to analyze and simulate the process and find the optimal and most efficient positionsof the panels in order to obtain the maximum amount of electricity and the most efficient thermodynamic characteristics of the system. Fig.6. and Fig.7. represents 3D model of the experimental setup of PV/T system. Experimental setup will be placed in front of the building of the Faculty of Mechanical Engineering Banja Luka. Whole setup will be set in spring 2020 when will start experimental research and all measurmments.



Fig.6. Front view of the experimental setup of PV/T system

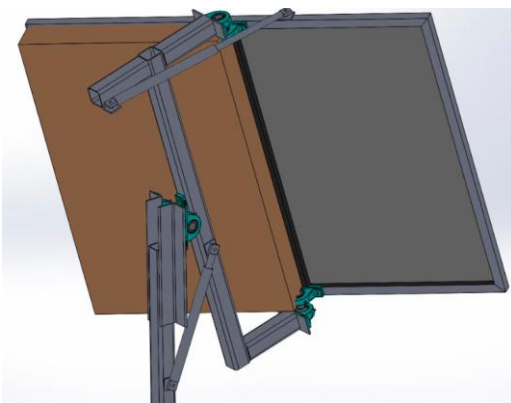


Fig.7. Back view of the experimental setup of TV/T system

Example of experimental setup of fix PV/T system is shown on Fig.8.

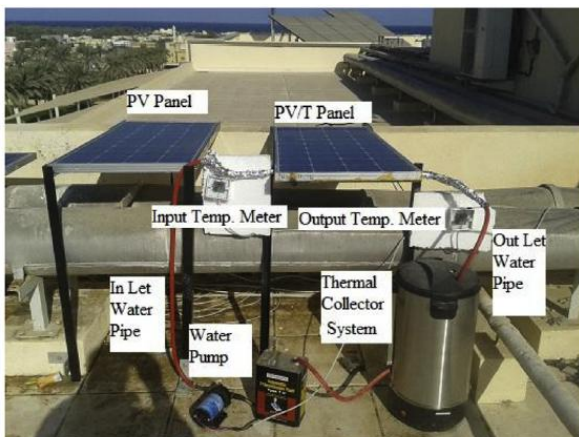


Fig.8. Experimental setup of PV/T system and referent PV panel[18]

4. Conclusion

PV/T systems are systems that utilize a wider range of solar radiation while generating electricity and heat. The application of these technologies is particularly interesting in urban areas where energy requirements are high and space for the installation of solar collectors and PV panels is limited. Integration of these systems and their use in the future represents a great potential for sustainable urban development. Combining the two technologies into one, a more efficient, less expensive and more convenient system is obtained in terms of installation and maintenance.

A large number of factors affecting the operation and efficiency of the PV/T system as well as the amount of electricity and heat generated, require additional research and analysis. Due to that, scientific project "Thermodynamic Analysis and Mechatronic Synthesis of Solar Power Plants in Urban Areas" implemented by Faculty of Mechanical Engineering in Banja Luka will contribute to this goal.

3. References

- [1] Abdelrazika, Al-Sulaimana, Saidurb, Ben-Mansoura, „A review on recent development for the design and packaging of hybrid,“ *Renewable and Sustainable Energy Reviews*, br. 95, pp. 110-129, 2018.
- [2] Browne, Norton, McCormack, „Heat retention of photovoltaic/thermal collector with PCM,“ *Solar Energy*, t. 133, pp. 533-548, 2016.
- [3] Sun, Li, Yuan, Cao, Lei, Yu, „Effect of tilt angle and connection mode of PVT modules on the energy efficiency of a hot water system for high-rise residential buildings,“ *Renewable Energy*, br. 93, pp. 291-301, 2016.
- [4] Sayyah, Horenstein, Mazumder, „Energy yield loss caused by dust deposition on photovoltaic panels,“ *Solar Energy*, br. 107, pp. 576-604, 2014.
- [5] Jaszczur, Teneta, Styszko, Hassan, Burzyńska, Marcinek, Łopian, „The field experiments and model of the natural dust deposition effects on photovoltaic module efficiency,“ *Environmental Science and Pollution Research*, t. 26, br. 19, pp. 8402-8419, 2019.
- [6] Gallardo-Saavedra, Karlsson, „Simulation, validation and analysis of shading effects on a PV system,“ *Solar Energy*, t. 170, pp. 828-839, 2018.
- [7] Tina, Cazzaniga, Rosa-Clot, Lot, „GEOGRAPHIC AND TECHNICAL FLOATING PHOTOVOLTAIC POTENTIAL,“ *Thermal Science*, t. 22, br. 3, pp. S831-S841, 2018.
- [8] Yushchenko, Bonob, Chatenoux, Patel, Ray, „GIS-based assessment of photovoltaic (PV) and concentrated solar power (CSP) generation potential in West Africa,“ *Renewable and Sustainable Energy Reviews*, t. 81, br. 2, pp. 2088-2103, 2018.
- [9] Bocca, Bergamasco, Fasano, Bottaccioli, Chiavazzo, Macii, Asinari, „Multiple-Regression Method for Fast Estimation of Solar Irradiation and Photovoltaic Energy Potentials over Europe and Africa,“ *Energies*, t. 11, 2018.
- [10] Katsumata, Nakada, Minemoto, Takakura, „Estimation of irradiance and outdoor performance of PV modules by meteorological data,“ *Solar Energy Materials and Solar Cells*, t. 95, pp. 199-202, 2011.
- [11] Nižetić. Čabo, Marinić-Kragić, Papadopoulos, „Experimental and numerical investigation of a backside convective cooling mechanism on photovoltaic panels,“ *Energy*, t. 111, br. C, pp. 211-225, 2016.
- [12] Odeh, Behnia, „Improving Photovoltaic Module Efficiency Using Water Cooling,“ *Heat Transfer Engineering*, t. 30, br. 6, pp. 499-505, 2010.
- [13] Chow, „A review on photovoltaic/thermal hybrid solar technology,“ *Applied Energy*, t. 87, pp. 365-379, 2010.
- [14] Kumar, Baredar, Qureshi, „Historical and recent development of photovoltaic thermal (PVT),“ *Renewable and Sustainable Energy Reviews*, t. 42, p. 1428-1436, 2015.
- [15] S. B. Debbarma, „Comparison of BIPV and BIPVT: A review,“ *Resource-Efficient Technologies*, t. 3, pp. 263-271, 2017.
- [16] Chow, He, Ji, „An experimental study of façade-integrated photovoltaic/water-heating system,“ *Applied Thermal Engineering*, t. 27, pp. 37-45, 2007.
- [17] Pola, Chianese, Bernasconi, „Flat roof integration of a-Si triple junction modules laminated together with flexible polyolefin membranes,“ *Solar Energy*, t. 81, pp. 1144-1158, 2007.
- [18] Kazem, „Evaluation and analysis of water-based photovoltaic/thermal (PV/T) system,“ *Case Studies in Thermal Engineering*, t. 13, 2019.

Information risk management in SME sector enterprises

Paweł Kobis

Czestochowa University of Technology, Faculty of Management, Częstochowa, Poland

pawel.kobis@wz.pcz.pl

ORCID: 0000-0003-0714-1888

Abstract: The article attempts to determine the level of use of methods increasing the security of information resources among the SME sector enterprises declaring the use of information security risk management methods. Research was carried out to determine the scope of use of risk management methods in the aspect of actions taken in the area of security of the intangible assets. Also the so-called "human factor" in the information protection process was taken into account. An attempt was made to determine how business entities use risk assessment in any form and how many of them use (and to what extent) the recommendations described in the ISO/IEC 27005 standard.

Keywords: INFORMATION, SECURITY MANAGEMENT, RISK, SME ENTERPRISES

1. Introduction

The majority of enterprises operating on the market process information in digital form. Data stored in paper form currently are a negligible part of all processed information and constitute rather an "analog copy" of a digital record. In addition, information is the company's most valuable resource. Processing information in a digital form involves the risk of its loss, destruction or takeover by third parties. This is primarily due to the fact that virtually every information system of a business entity operates in a local network with a direct connection to the Internet. Thus, potentially unsecured data and information resources can be accessed from virtually anywhere in the world. Therefore, it becomes necessary to develop appropriate principles of information management, implementation of specific safeguards, and identification, analysis and assessment of information risk. The risk management process is particularly important. Identifying the weaknesses of both the company's IT system and ways of securing digital content is a starting point in determining potential threats and the occurrence of unwanted incidents.

2. Theoretical outline of risk management in the area of information security

Risk management is the main area in the information security management system in business entities. This process can be divided into six main stages forming a closed loop [1,2]:

- Context Establishment;
- Risk assessment consisting of:
 - Risk identification;
 - Risk analysis;
 - Risk evaluation
- Risk treatment;
- Risk acceptance;
- Risk communication and consultation;
- Risk monitoring and review.

Closed loop proves that this process is being repeated successively in entities, and security management is continuous. Any, even the slightest change in the information system should generate the need to re-perform the risk management procedure. A detailed theoretical description of the individual stages of risk management together with the methods currently used in organizations for estimating the appropriate quantities could successfully constitute material for a separate book publication. In this article, in the theoretical part, only definitions of elements of this process are provided. The stages of risk management are graphically presented in Figure 1. This is a model contained in the international standard ISO/IEC 27005 [3].

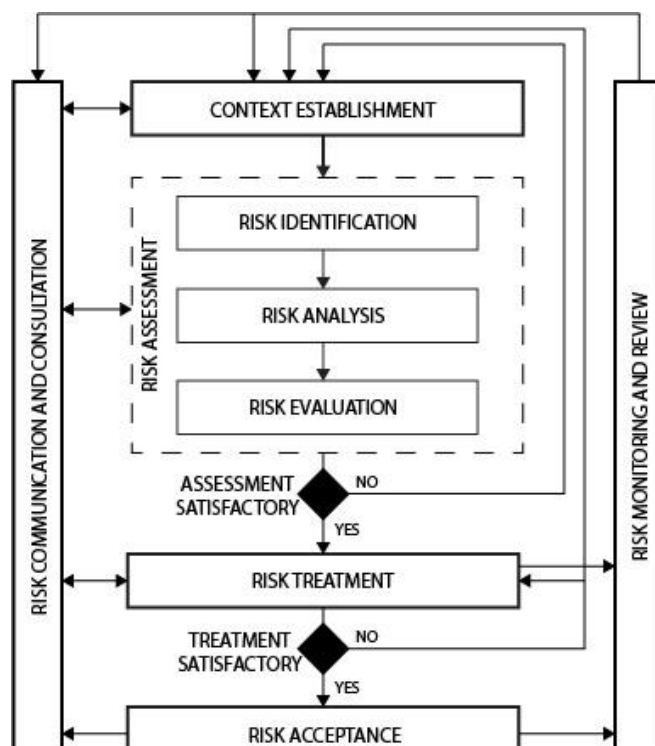


Fig. 1 Stages of risk management in the field of information security in business entities.

Source: [4]

These stages are a logical sequence of actions aimed at minimizing or, in exceptional cases, eliminating potential risks affecting the security of digital resources [5]. They "recognize" potential information threats, allow you to accurately determine the structure of your own information system - identify its strengths and weaknesses. They allow the company to be identified as part of a wider, regional, national and perhaps even global information exchange system. The individual stages of risk management can be briefly characterized as follows:

1. Context Establishment - involves analysing the business and organizational environment of the enterprise. It consists in a detailed separation of external and internal factors affecting information processing in the enterprise [6]. In the case of external factors, the most common ones include: competition of the business entity, relations with clients and business partners, political and cultural environment, technological environment in the aspect of the possibility of using the existing potential [7]. The legal environment and related rights and obligations of the organization associated with the processing of information resources are also important. In the case of internal factors, we are talking about: organization of information processing within the enterprise, organizational structure, strategy and goals of the economic entity, the potential of the organization (technical, resource, knowledge, etc. potential).

2. Risk assessment - a key element of the information security management system. The first stage is to determine the causes and ways of materializing unwanted incidents. As part of risk identification, it is necessary to list all database resources, all security measures used and usable and to list all possible threats and vulnerabilities [8]. An inventory of assets should be made, in particular: software, hardware, technologies used, locations and people. It is very important to identify individual thematic areas in the field of information processing, specification of groups, teams, people who have access to information areas in a specific way and scope. It is important here to take into account the human factor, i.e. the human impact on information security - identification of strengths and weaknesses in the aspect of information security. Thus, it is necessary to determine the level of knowledge and experience of individual employees. This is a difficult task, but necessary in the times of continuously emerging social engineering techniques. In terms of the human factor, it is also necessary to specify the level of possibilities for conducting possible trainings for employees of the business entity - training facilities. Many techniques [9] are used to identify risk. The most important of them include [10]:

- review of the documentation;
- information gathering techniques, which include: brainstorming, Delphi technique, surveys and root cause analysis (RCA);
- checklists based on an analysis of similar projects from the past;
- analysis of project assumptions;
- techniques based on diagrams: cause and effect diagram, impact diagrams, block diagrams;
- SWOT analysis;
- expert assessment.

The second stage of risk assessment is its analysis. It allows with a certain approximation or probability to estimate the occurrence of a threat and its impact on information security.

Many methods of risk analysis are used in management theory and practice (CRAMM, COBRA, MARION, FMEA, OCTAVE, MEHARI, ISACA and others) [11, 12]. Many business entities develop their own proprietary methods. Generally, they can be divided into 3 groups [13]:

- Quantitative methods - in which attempts are made to quantify the probability of their occurrence and potential losses;
- Qualitative methods - based on the assessment of threats, their significance and possible losses associated with them, based on good practices and experience of a person or group of assessors;
- Mixed methods - using elements of the above two methods.

The third stage is risk assessment. It allows us to answer the question: is the risk acceptable? Whether any consequences of its occurrence constitute, or not, serious problems for the enterprise. At this stage, the risks are prioritized and in case of their occurrence the priorities for action are defined. The so-called residual risks are also specified, i.e. risks that remain after all possible or economically reasonable steps have been undertaken to avoid those risks. Risk assessment is an introduction to the stage of risk management.

3. Risk management means a specific set of actions to be undertaken in relation to the estimated risks [14]. It aims to provide certain options that will reduce or eliminate the risks if they occur. Risk management should take into account the effects of the risk (be proportionate to them) and should be cost-effective. The most common risk procedures include:

- Risk avoidance - implemented through resigning from certain actions that cause risks, replacing some actions with other, less risky ones, etc.
- Risk transfer - consisting in transferring responsibility for the effects of risk to another entity. This includes, for example, insurance, guarantees, employment of

subcontractors (separate business entities) for selected risk-taking activities.

- Risk mitigation - a set of methods that reduce the likelihood of risk occurrence to an acceptable level.
 - Risk acceptance - means the fact of knowingly accepting the risk and not taking any action to reduce it. It occurs most often in case of the absence of appropriate countermeasures despite access to all possible information related to the subject. It should be noted that accepting risk without first thoroughly investigating the phenomenon with as much information as possible can be considered as ignoring risk. We distinguish active and passive acceptance of risk. Active acceptance assumes creating a contingency plan in the event of a risk. In the case of passive acceptance, there is no contingency plan and no actions are taken [10].
4. Risk acceptance is the stage at which decision makers accept the risks developed in previous points of risks, their possible effects and established actions. Dealing with risk seems to be satisfactory [15].
 5. The risk communication and consultation stage is a kind of stage covering all 4 previous stages. It runs in parallel and allows communication between decision-makers and other interested parties within the scope of the exchange of information on the risk and reaching agreement on the risk management [16].
 6. The risk monitoring and review stage is a continuous process of tracking the risks and effectiveness of the security measures implemented by the organization [17]. It involves assessing the procedures developed as part of risk management and estimating their timeliness. It enables the analysis of the state of the information system - its possible changes, needs for updating and, as a consequence, conducting a new risk management procedure, taking into account all its stages.

The risk management scheme according to ISO 27005 is quite flexible and allows decision-makers of business entities to repeat any given stage of the process. At the same time, it is consistent in their implementation, not giving the opportunity to bypass any of them [18]. Thus, the authors of the scheme indicate that risk management is not a "rigid" process. Its course largely depends on the business entity, its needs and experience in the field of risk management. The authors of the scheme and other provisions of the standard also do not divide business organizations depending on risk management. Thus, specific guidelines are equally valid for small, medium and large enterprises. Certain 'framework' of proceedings is identical, however: the scope of analysed issues and the way of functioning of information systems in enterprises are changing.

3. Research results and discussion

Research using the CAWI (Computer-Assisted Web Interview) method was conducted in the period from January 2019 to June 2019 in 117 enterprises of the SME sector in the Silesian Voivodeship in Poland. The selection of enterprises for research was made using the purposive sampling method, taking into account the following condition: the business entities carried out to any extent the risk analysis of information security. People directly involved in business organizations in information security processes were asked to complete the survey. The survey was placed on the website of the author of this article. Requests for completing the survey were sent by e-mail to companies' email addresses and were communicated directly by phone.

The first question (Table 1) concerned the definition of resources and areas taken into account when assessing risk in business entities. The question includes traditional areas of information processing as well as those related to the human factor. All respondents declared here the area of hardware resources - thus

recognizing them as the most important in the information management system. Most of the respondents (98%, 99%, 98% and 99% respectively) also included: software resources, database resources, used security measures, and potential threats. Slightly less, because 87% also include technologies used in risk assessment. The above areas constitute a classic canon in the probed business entities in determining potential risk. Only slightly more than half of enterprises identify vulnerabilities in their information systems. This is quite surprising, since the vulnerability area should be one of the starting points when assessing risk. It follows that the potential effects of a threat are determined without establishing a precise genesis of the possibility of their occurrence. This is a methodological error showing the fundamental lack of knowledge among those responsible for the protection of intangible resources.

Table 1 Answers of the respondents to the question: Please indicate which areas of risk identification do you include in the process of risk estimation

1	hardware resources	100%
2	software resources	98%
3	database resources;	99%
4	locations of information resources	26%
5	used security measures	98%
6	technologies used	87%
7	potential threats	99%
8	identified vulnerabilities	54%
9	level of training	14%
10	employee susceptibility to social engineering methods	29%
11	employee knowledge	38%
12	employee experience	18%

Source: own study

Most business entities surveyed do not take human factors into account when estimating risk [19]. The highest result obtained in this group of areas is employee knowledge of 38% and social engineering methods with a result of 29%. Employee knowledge significantly affects information security. Knowledge of new techniques and technologies of protection as well as used attack techniques allows for an appropriate response from the employee. In addition, having knowledge about information security, the employee properly organizes the workplace (also in the virtual space) so as to minimize the risk of any incident. Therefore, it is surprising that only 38% of respondents pay attention to this. The training aspect is completely negligible (only 14% of responses). The pace of development of information management methods in digital form and the pace of development of more and more new techniques of digital aggression force continuous expansion of knowledge in this area. This applies not only to people who are formally responsible for security, but to any employee who processes any information. Therefore, when estimating risk, special attention should be paid to e.g. the planned frequency of training - this will reduce the likelihood of multiple incidents.

The next question concerned methods used by company representatives when assessing the risk of information security breach (Fig. 2).

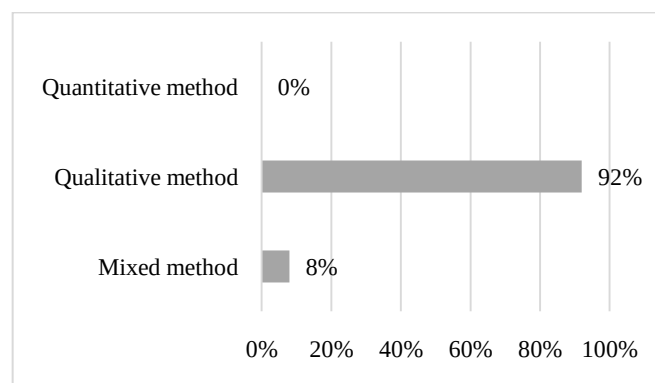


Fig. 2 Answers of the respondents to the question: Please specify which methods of risk assessment do you use in your company?

Source: own study

Most respondents indicated qualitative methods. Only 8% of the surveyed entities use mixed methods and none of them uses only quantitative methods. Where do these proportions come from? This probably results first of all from the specificity of the area in which the risk is determined, and secondly from the level of difficulty of individual methods. The area of information and the associated risk of security breach and, consequently, the area of resulting effects is sometimes difficult to quantify. Example: how to quantify the level of loss of reputation among customers of a business entity after occurrence of a specific incident? It seems to be very difficult; or: how to quantify the level of confidence of business partners in an enterprise that has a given information protection system. There are more examples of this type. That is why economic organizations very often lean towards qualitative methods in which (to simplify) the level is expressed subjectively on a descriptive scale (low, medium, high). The benefits of qualitative methods are [20] as follows:

- no need to value information (its availability, confidentiality, integrity) - it is very difficult and sometimes impossible - the value of information is a relative and variable concept;
- no need to estimate the costs of recommended risk management methods or calculate potential profit / loss;
- the possibility to identify general, significant risk areas;
- the possibility to consider and take into account, when estimating risk, the aspects that are very difficult to measure, such as company image, organizational culture, etc.;

Qualitative methods are more understandable to most people - they do not contain a complicated mathematical apparatus. For example, considering the answers in Figure 3, almost half of the respondents carried out a risk analysis only because of a certain requirement dictated by the introduction of the provisions of the GDPR. As much as 26% also perform the analysis less than once a year. Therefore, it should be assumed that all these business entities take up decisions guided by the choice of less work effort.

Going back, however, to risk assessment methods, despite the fact that none of the entities declared to use only quantitative methods, 8% declared mixed methods in which there are elements of methods containing a mathematical apparatus. Quantitative methods bring several benefits [20]: estimation and results are objective and comparable, the value of information is expressed in money and the results of risk estimation are expressed in to dimensions: financial and percentage. It should be assumed that 8% of business entities have more objective risk assessment results than the rest.

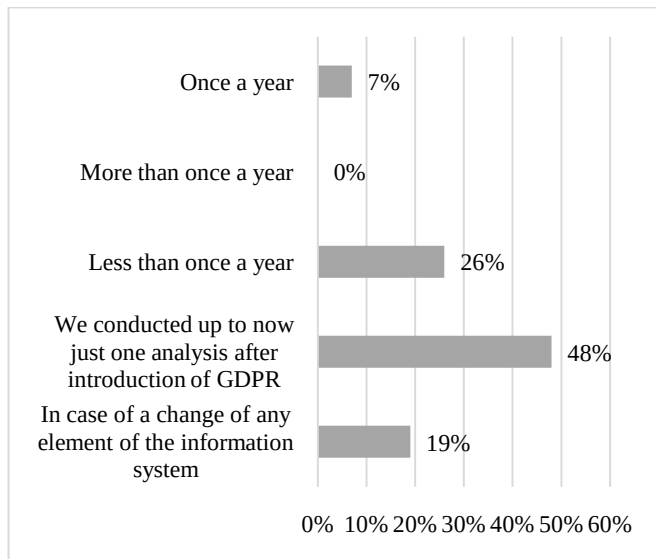


Fig. 3 Answers of the respondents to the question: How often do you carry out the various stages of the risk management process
Source: own study

As mentioned in the theoretical part of the study, risk management processes in the aspect of information security should be carried out in full whenever any element of the information system is changed. In the case of the surveyed business entities, only 19% declare this type of practice (Fig. 3). This may be due to the fact that only 28% of the surveyed enterprises take into account the recommendations of the ISO/IEC 27005 standard and 11% only part of the recommendations described in that standard (Fig. 4).

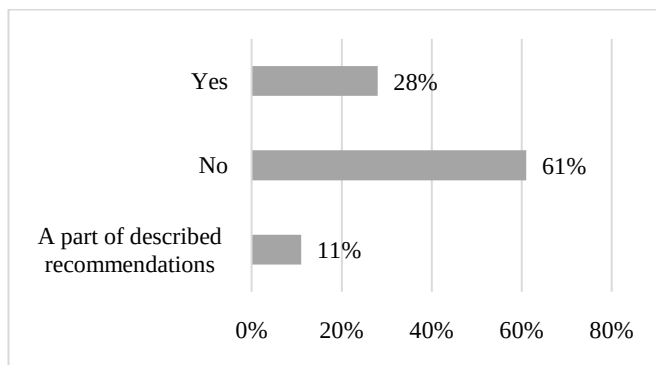


Fig. 4 Answers of the respondents to the question: Do you include ISO/IEC 27005 recommendations in the information security risk management process?
Source: own study

The majority of respondents (61%) admit that they do not apply the recommendations of an international standard during risk management. In a sense, this coincides with the answer to the question about the frequency of risk management. Entities that implement these procedures only because of the entry of the GDPR probably use ready-made (template) solutions offered by commercial companies for risk analysis. This, of course, does not mean that the lack of consideration of the ISO standard is equivalent to poor quality of the analyses. However, it should be assumed that they may slightly differ from standards generally accepted in the world.

The next, and last question asked to the respondents concerned the awareness of the respondents about the impact of the human factor on the security of information resources (Table 2).

Table 2. Respondents replied to the question: Please specify to what extent (1- smallest, 5-largest) in your enterprise the information security may be influenced by the following factors

1	social engineering activities	1.9
2	lack of knowledge	3.8
3	lack of experience	3.2
4	indifferent attitude to the work performed	1.6
5	fatigue	1.4
6	intentional, harmful actions of employees	1.1

Source: own study

Persons responsible for information security were asked to estimate the likelihood of factors having a negative impact on intangible assets directly related to employees. Thus, an attempt was made to answer the question: is the human factor in business entities treated equally seriously with other threats? According to respondents, the biggest threat is the lack of appropriate knowledge and experience among employees. This is the right approach - taken into account in 38% of process in the area of risk identification (Table 1). This may not be a highly satisfactory result, but quite optimistic compared to the other answers. In the adopted Likert scale from 1 to 5, the respondents estimated below the value of 2 the impact on information security of such factors as social engineering activities, indifferent attitudes to their work, fatigue or deliberate, harmful activities of employees. This is a surprising result, because IT security specialists officially confirm that cybercriminals use fraudulent tricks on a large scale to gain access to confidential company information and manipulate employees, and the human factor is the biggest challenge for companies in ensuring the expected level of security [21]. So how to justify the answers received? - perhaps respondents speaking about their own companies do not believe that these factors occur in their companies and employees are properly prepared in this aspect. But if one adopts this way of reasoning, the question should be asked: on what basis do they make such assumptions, since training in such a dynamically changing area of knowledge is marginalized at the stage of risk assessment (14% - Table 1).

4. Conclusions

Information security is currently one of the most dynamically developing areas of information management. Risk management is the crucial from the point of view of protecting intangible assets. It allows, as a result of the conducted identifications, analyses and assessments, to implement an effective information security policy in a business entity. Risk management is a process and as such it requires special diligence at every stage carried out. Each stage is the result of the previous one and contributes to the next one. That is why it is very important to follow specific guidelines described in detail in international standards: the guidelines to guarantee the best possible results.

This article includes research conducted in 117 business entities of the SME sector. The results obtained outline the overall picture of how the risk management process works in practice. Both strengths and weaknesses of this process are highlighted. Little interest of enterprises in the human factor in risk assessment and in identifying potential dangers caused by employees has been revealed. The research results encourage further observation of enterprises in the field of risk management, which will allow in the future to estimate the rate of growth of the awareness of business entities within the scope of use of recommendations of international standards.

5. References

1. M. Ghazouani, H. Medromi, A. Sayouti, S. Faris, *Information Security Risk Assessment - A Practical Approach with a Mathematical Formulation of Risk*, International Journal of Computer Applications, **103**(8), 36–42, (2014).
2. T. Peng, *ISO 27005 Information Security Risk Management* [Online]. SlideShare, (2016), [access date: 14.01.2020], URL: <https://pt.slideshare.net/thomaspeng30/iso-27005-information-security-risk-management-download-free-template/7>
3. V. Agrawal, *A Framework for the Information Classification in ISO 27005 Standard*, 2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud), 264–269, (2017).
4. *ISO/IEC 27005:2014-01 Standard*, [Online], 2014, [access date: 14.01.2020], <https://sklep.pkn.pl/pn-iso-iec-27005-2014-01p.html>
5. B. Barafort, A.L. Mesquida, A. Mas, *Integrating risk management in IT settings from ISO standards and management systems perspectives*, Computer Standards & Interfaces, **54**, 176–185, (2017).
6. M.A. Fikri, F.A. Putra, Y. Suryanto, K. Ramli, *Risk Assessment Using NIST SP 800-30 Revision 1 and ISO 27005 Combination Technique in Profit-Based Organization: Case Study of ZZZ Information System Application in ABC Agency*, Procedia Computer Science, **161**, 1206–1215, (2019).
7. C. Yang-Ngam, T. Chankoson, P. Aodton, *Influence of internal and external factors on supply chain information system risk management implementation*, International Journal of Supply Chain Management, **8**, 612–623, (2019).
8. A. Shamel-Sendi, R. Aghababaei-Barzegar, M. Cheriet, *Taxonomy of information security risk assessment (ISRA)*, Computers & Security, **57**, 14–30, (2016).
9. L. Pan, A. Tomlinson, *A systematic review of information security risk assessment*, International Journal of Safety and Security Engineering, **6**, 270–281, (2016).
10. K. Zielosko, *Analiza ryzyka*, [Online], Encyklopedia Zarządzania, (2019), [access date: 14.01.2020], URL: https://mfiles.pl/pl/index.php/Analiza_ryzyka
11. J. Łuczak, *Metody szacowania ryzyka – kluczowy element systemu zarządzania bezpieczeństwem informacji ISO/IEC 27001*, Zeszyty Naukowe Akademii Morskiej w Szczecinie, **19**, 63–70, (2009).
12. G. Wangen, *Information Security Risk Assessment: A Method Comparison*, Computer, **50**(4), 52–61, (2017).
13. J. Stanik, M. Kiedrowicz, *Metoda analizy i szacowania ryzyka zasobu informacyjnego*, Roczniki Kolegium Analiz Ekonomicznych, Szkoła Główna Handlowa, Warszawa, **49**, 371–390 (2018).
14. K. Mersinas, B. Hartig, K. Martin, A. Seltzer, *Measuring Attitude towards Risk Treatment Actions amongst Information Security Professionals: An Experimental Approach*, Conference: Workshop on the Economics of Information Security, At Berkeley, CA, (2016).
15. S. Snedaker, C. Rima, *Chapter 6 - Risk Mitigation Strategy Development*, [in:] S. Snedaker, C. Rima (Eds.), *Business Continuity and Disaster Recovery Planning for IT Professionals*, Syngress, 337–367, (2014).
16. R.J. Chapman, *Simple Tools and Techniques for Enterprise Risk Management*, John Wiley & Sons, New Jersey, (2011).
17. C. Martani, *Risk Management in Architectural Design: Control of Uncertainty over Building Use and Maintenance*, Springer, Cham, (2014).
18. S. Ariyani, M. Sudarma, *Implementation Of The ISO/IEC 27005 In Risk Security Analysis Of Management Information System*, Journal of Engineering Research and Application, **6**(8), pp.01-06, (2016).
19. P. Kobis, *Human factor in the aspect of digital information in business enterprises*, Proceedings of the 9th International Conference on Management (ICOM) vol. II, Gödöllő, Hungary, 35–42, (13-14 June 2019).
20. P. Kawczyński, *Analiza ryzyka – metody szacowania ryzyka – cz. 2*, [Online], PortalODO by Lubasz i wspólnicy, (2014), [access date: 15.01.2020], URL: <https://portalodo.com/analiza-ryzyka-metody-szacowania-ryzyka-cz-2/>
21. KPMG, (2019). Report: *Barometr cyberbezpieczeństwa. W obronie przed cyberatakami*, [Online], KPMG, (2019), [access date: 17.01.2020], URL: <https://assets.kpmg/content/dam/kpmg/pl/pdf/2019/04/pl-Raport-KPMG-Barometr-Cyberbezpieczenstwa-W-obronie-przed-cyberatakami.pdf>

Introduction of Industry 4.0 in Industrial Enterprises: Problems and Challenges

Neli Veleva¹

Faculty of Mechanical Engineering, Department of Industrial Management – Technical University of Varna, Bulgaria¹

neli_veleva@tu-varna.bg

Abstract: This publication analyzes the opportunities, problems and challenges associated with the implementation of Industry 4.0 in Bulgarian industrial enterprises in condition of a dynamic changing competitive environment and extending digitalization. An analysis is made on the base of the four fundamental principles that are used for application of solutions related to the implementation of Industry 4.0 and that influence on the organizational activities in all functional areas in the industrial enterprise. Various industrial sectors are analyzed to identify the trends, problems and challenges of the introduction of Industry 4.0 in manufacturing activities. Opportunities for improving their competitiveness are identified by means of using the principles and advantages of the modern digitalization. Up-to-date statistics and various methods for analyzing of the existing information are used. The specifics of development of the leading sectors in Bulgarian economy are identified. The results show that there is a growing number of Bulgarian industrial enterprises focused on deploying Industry 4.0 to reach new higher levels of productivity, quality and efficiency. This is a new approach in seeking favorable opportunities for their further sustainable development and competitiveness. Some fundamental problems, challenges and barriers in industrial enterprises, in the process of introduction of digitalization in the industry are discussed, which hinder the fast pace of their development. As a result of this analysis are presented some appropriate solutions.

Keywords: INDUSTRIAL PRODUCTION, INDUSTRY 4.0, DIGITALIZATION, CHALLENGES, SUSTAINABLE DEVELOPMENT.

1. Въведение

In the modern world, it is increasingly being talked about the enormous impact that digitalization is developing at an extraordinary rate, its advantages and challenges. The process itself may be called "megatrend", which causes the world to look toward new, more sophisticated levels of productivity. More and more organizations are redirecting and refining their strategies and goals toward the introduction of new generations of digital technology. Aiming at building competitiveness, today every industrial enterprise is concentrating its efforts and resources on the development and improvement of digital technologies through the implementation of Industry 4.0. The aim is to achieve higher levels of productivity, quality, efficiency, search for favorable opportunities for their further sustainable development and competitiveness.

2. Industry 4.0 - essence and characteristics

Industry 4.0 is a modern phenomenon that has many advantages but also has negative impacts. There is no one specific answer for the name "For" or "Against" Industry 4.0. Prior the introduction of Industry 4.0, every industrial enterprise must clarify the specific nature and specific characteristics contained in the concept of digitalization through its prism of reality. According to the concept of digital transformation of Bulgarian industry, "... the importance of new technologies for society and economy is determined by the definition of the term "Fourth Industrial Revolution"[2]. Based on the detailed literature review, could be done the conclusion about the majority of the Fourth Industrial Revolution. These conclusions may be presented by summaries of researchers in this field (Industry 4.0) (Table 1):

The Fourth Industrial Revolution is developing on the base of modern digitalization, combining different technologies, methods and digital solutions, which lead to significant changes in the market environment, business and society.

Table 1: Key summaries concerning the essence of the Fourth Industrial Revolution (Industry 4.0).

	The Fourth Industrial Revolution (Industry 4.0) - The Essence
Federal Government of Germany, 2011 (High-Tech Strategy for Germany 2020)	„High-tech strategy for the development of the German industry in 2011“ [3]
According to the Concept for Digital Transformation of Bulgarian Industry	"Industry 4.0. is defined as part of the application of new digital technologies in the manufacturing sector and includes a wide range of technological solutions and business models

[2017]	that contribute to qualitatively new forms of economic activity. "[2]
Corresponding Member MD Svetozar Margenov Director of the Institute of Information and Communication Technologies - BAS (IICT - BAS),	"Digitization in industry: ecosystem for sustainable innovative growth "[5]
Prof. Dr. J. Ovcharova (Head of the Institute of Information Technology Management (IMI))	„The term Industry 4.0 is a word game consisting of 3 parts - first "industry" means that this initiative is aimed at the industry; second, "0" is a reference to Internet technologies; and third "4" marks the fourth industrial revolution and aims to reach a new level of industrial value ... "[6]

The need to understand, use and fully implement the concept of Industry 4.0 is primarily due to the fast development of new technologies leading to the automation, modernization, refinement and digitization of the real production and business processes. In this way, each organization (industrial enterprise) manages to achieve sustainable development and growth and high levels of competitiveness.

According to the Concept for Digital Transformation of Bulgarian Industry from 2017 digitalization of the economy "..... it is developing dynamically and is an important driver of innovation, competitiveness and growth". The Fourth Industrial Revolution is based on four basic principles that are used to implement decisions related to its implementation and that affect organizational activities in all functional areas of an industrial enterprise. These are: [2]

1. **Interoperability** - the ability and need of a system or product to work with other systems or products without the need for additional user action. This principle is implemented through the Internet of Things (IoT) or Internet of People (IoP);

2. **Transparency** – it is realized by the ability of new generations of information systems to calculate and store data in the so-called cloud structures.

3. **Structuring and visualizing information** in a new, comprehensible way that helps managers to make decisions about pressing issues.

4. **Autonomy** - technological innovations are not universally applicable, but they must be tailored to the specific functional areas of the industrial enterprise.

Applying and complying these 4 fundamental principles is a key factor in achieving the competitiveness, development, productivity and survival of any organization in today's competitive environment. [10]

Industry 4.0 is a set of various activities, business models, technological and strategic decisions. E-commerce, intelligent manufacturing enterprises, cloud technologies, internet technologies, artificial intelligence, 3D printing, robotics, smart cities, smart factories are only a small part of the development of automation and real-time data exchange in manufacturing processes.

The characteristics of the Fourth Technological Revolution may be summarized as follows:

- Decision-making optimization - Big Data analysis and processing facilitates more flexible decisions when urgent problems arise;
- Resource Productivity and Resource Efficiency - The pursuit of basic strategic goals, namely producing the most output based on available resources at the lowest possible cost. In this way, optimization of production processes, reduction of costs, better use of resources and energy, reduction of harmful emissions from production and environmental protection are achieved.
- Customer/customer feedback regarding ordering, planning, manufacturing, design regarding the production of a product or service.
- Dynamic organization of production processes according to time, quality, price, risk, sustainability, resources, suppliers, customers and more.

3. Introducing Industry 4.0 into Industrial Enterprises - Essence, Problems, and Challenges

"The Industry 4.0 platform is a response to the changes in modern technology that have led to the fourth industrial revolution since the introduction of machines, production lines and IT." [1]

The introduction and implementation of this platform is not an easy process for any industrial enterprise, which is developing in the current competitive environment. In an extremely fast-paced, dynamic business environment, it is clear that there is no one-size-fits-all solution or sequence that can be applied to digitalization, both for large, small, and medium-sized industrial enterprises. Each organization is unique and each decision must be strictly individual. The desire to achieve pragmatism and adaptability on the part of enterprises is increasingly being sought, trying to achieve maximum results with limited resources and opportunities.

One of the main keys to the introduction of Industry 4.0 is the promotion of human-machine collaboration. The pace of introduction of new technologies, resources and staff qualification are also among the key factors for transition to the Fourth Industrial Revolution.

According to the Strategy for Bulgaria's Participation in the Fourth Industrial Revolution. (Industry 4.0) – 2018, [9] 4 basic strategic tasks for the introduction of Industry 4.0 can be identified and practically implemented as they are listed:

- Institutional support for the development of Industry 4.0 - identification of specific strategies, policies, programs, procedures, mechanisms and tools to support the development of Industry 4.0 in relation to the peculiarities of Bulgarian organizations and market conditions;
- Creating synergies between existing policies, programs, procedures, strategies and support mechanisms;
- Development of educational and scientific initiatives to build institutional and organizational capacity for acceptance and compliance with the requirements and prerequisites of Industry 4.0;
- Creating preliminary projects and demonstration models for visualization, analysis and implementation;

Digital education, qualifications and training in the range of the work process can be cited as major problems with implementation. Their solutions lie in providing accessible tools for employees that are used in different departments of industrial enterprises. In this way, by sharing experience and information, organizations can achieve their ultimate goals faster, pragmatically, by joining forces to implement complex solutions.

The key to the introduction of Industry 4.0 is based on the seamlessly digital knowledge transformation into everyday skills and professional qualifications - fast, pragmatic and more active. [10]

The significant role of innovation must also be noted here. To realize an innovative idea, several components are needed: creativity, entrepreneurship, freedom. [1]

In the classification of the main challenges and obstacles to the introduction of Industry 4.0 Industrial enterprises can be identified the following risks and barriers:

- Unclear economic benefits;
- Investment amount;
- Qualification and training of employees;
- Insufficient standardization and certification;
- Unclear economic conditions;
- Data security;
- Strong competition;
- Insufficient stimulation of digitalization by the state;
- Adaptation of legislation;

Some of the major benefits that any industrial enterprise could benefit after deciding to introduce and use the latest generation digital platform are:

- Increasing efficiency;
- Optimization of planning and management processes;
- Achieving resource efficiency and resource productivity;
- Maintaining competitors advantages and competitiveness in the market;
- Collection and analysis of large amounts of data through information systems;
- Cost reduction;
- Increase in turnover;
- Achieving sustainable development and growth;
- Employee-user feedback;

According to the National Statistical Institute (NSI) the results of the conducted in 2019 a survey of enterprises with 10 or more employees show that 95.7% of them use computers, and for large enterprises with 250 and more employees the relative share is 100.0%. High-speed, reliable and uninterrupted Internet access is a fundamental necessity for businesses and 93.7% of them have permanent access to the global network. To connect with Internet companies, they mainly use DSL or other fixed technology (80.5%), with 64.8% of them having the fastest fixed connection speed exceeding 30 Mbps. [11]

From the Fig. 1, it is clear that more and more organizations are using the Internet, which is directly dependent on the development and implementation of Industry 4.0 in the Bulgarian market economy. [9]

Another important factor concerning the essence of the introduction of Industry 4.0 is the provision, deployment and use of information and communication technology (ICT) in an industrial enterprise. [8]

According to NSI data, in 2019 one of five enterprises (20.1%) has employees whose primary job is to develop, manage or maintain ICT systems or software applications. In the previous 2018 10.0% of enterprises have employed or have tried to employ ICT specialists, as 41.6% of them have experienced difficulties in finding suitable staff. 10.1% of employers provide ICT training to improve the skills of their staff. Fig. 2 presents the relative share of

enterprises that employ ICT professionals and provide training in the field of ICT by enterprise size in 2019. [11]

Share of enterprises with internet access by size class and economic activity in 2019

	Internet	Fixed connection	From them: with speed ≥ 30Mbps
(Per cent)			
Total	93.7	80.5	64.8
By size class			
10 - 49 persons employed	92.5	78.3	62.3
50 - 249 persons employed	98.8	89.7	73.8
250 or more persons employed	100.0	97.5	81.8
By economic activity			
Manufacturing	94.1	81.7	59.5
Electricity, gas and steam, water supply, sewerage and waste management	98.9	84.7	69.0
Construction	95.2	78.1	66.6
Wholesale and retail trade; repair of motor vehicles and motorcycles	93.8	80.7	59.7
Transportation and storage	95.6	78.2	58.8
Accommodation and food service activities	85.4	70.1	67.0
Information and communication	100.0	96.0	93.8
Real estate activities	97.5	87.1	75.6
Professional, scientific and technical activities	99.7	94.0	79.6
Administrative and support activities	87.3	73.7	71.1

Fig. 1 Relative share of enterprises with Internet access of enterprises economic activities in 2019 by size according to NSI [11]

Share of enterprises employing ICT specialists and providing ICT training by size class in 2019

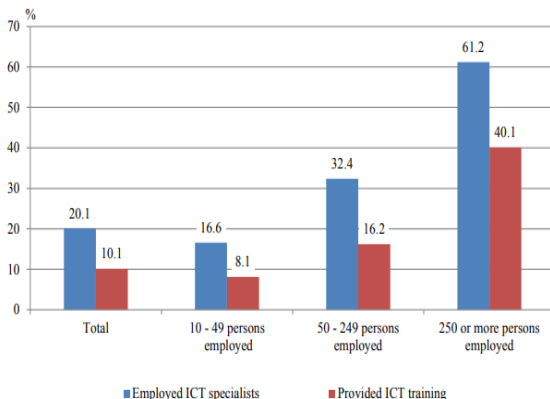


Fig. 2 Share of enterprises employing ICT specialists and providing ICT training by size class in 2019 according to NSI [11]

More and more Bulgarian industrial enterprises are focusing on the implementation of Industry 4.0 to achieve higher levels of productivity, quality and efficiency. This is a new, modern approach to seeking opportunities for their further sustainable development and competitiveness. [4], [7].

4. Barriers to industrial enterprises in the process of introduction of digitalization in industry and appropriate solutions.

The barriers defined by the analysis of existing information facing industrial enterprises determined to introduce and use new generations of technology can be classified into several major groups.

- Related to policy - One of the most important barriers facing industrial organizations is the need for long-term and sustainable national and international policies that have to ensure the development of smart innovation and sustainable change.
- Related to new technologies - in this group, the main barrier is that enterprises must be uncompromisingly quickly ready to apply and use new technologies, creating innovation. Flexibility of the organization regarding to this barrier is crucial in achieving competitive positions in the face of new market conditions.

- Related to R&D. Research and development (R&D) needs adequate financial support, both from the EU and from a national perspective.
- Related to qualification - the shortage of highly qualified personnel also influences the issue of introducing and embarking on the Fourth Technological Revolution.

5. Conclusion.

Achieving resource productivity (producing as much output as possible from available resources) and resource efficiency (at the lowest possible consumption of resources for available quantities of output), individual approach to the consumer, flexibility and potential for value creation by offering new services and products, summarizes and presents how important and necessary the process of deploying Industry 4.0 in an industrial enterprise is. In this way, the key advantages considered are the challenges and benefits of implementing digitalization, defining Industry 4.0 as a powerful tool for achieving high efficiency and competitiveness of industrial enterprises in a highly competitive environment.

This research was carried out in the range of research project NP14/2019 of Department of Industrial Management at the Technical University of Varna and the research work on this problematic will continue after the project completion.

6. References.

1. Borisov, B. Industry 4.0 or what digitalization means to the industry, Conference „Industry 4.0 – Challenges and Implications for Bulgaria's Economic and Social Development“ from 12th April 2018 in Sofia, Republic of Bulgaria;
2. Concept for digital transformation of the Bulgarian industry (Industry 4.0) 2017-2030;
3. High-Tech Strategy for Germany 2020
4. Krasimira Dimitrova, Technology, Quality and Innovation - Key Aspects of Competitiveness of Industrial Enterprises, International Conference Applied Computer Technologies (ACT) 2019 from 19th to 21th September in Varna, Republic of Bulgaria.
5. Martenov, Sv., Digitalization in the industry: an ecosystem for sustainable innovative growth - Conference „Industry 4.0 – Challenges and Implications for Bulgaria's Economic and Social Development“ from 12th April 2018 in Sofia, Republic of Bulgaria;
6. Ovcharova, Jivka, Implementation of Industry 4.0 in SMEs: Minimizing Risks and Identifying Opportunities –Conference „Industry 4.0 – Challenges and Implications for Bulgaria's Economic and Social Development“ from 12th April 2018 in Sofia, Republic of Bulgaria;
7. Panayotova, Tanya & Dimitrova, Krasimira (2019). A Strategic Vision for Development of Flexible Industrial Enterprise, Chapter 11 in DAAAM International Scientific Book 2019, pp.143-158, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-24-2, ISSN 1726-9687, Vienna, Austria;
8. Panayotova, Tanya, Opportunities for research the impact of intellectual capital on creating value in industrial enterprises, International Conference Applied Computer Technologies (ACT) 2019 from 19th to 21th September in Varna, Republic of Bulgaria;
9. Strategy for Bulgaria's Participation in the Fourth Industrial Revolution (Industry 4.0) of 2018.
10. Velikov, M. „Industry 4.0. Will replace the machines people?“ , Journal „your BUSINESS“, issue 3 , 2017
11. National Statistical Institute – www.nsi.bg

On predictability of precious metals towards robust trading

Turan Erman Erkan^{1,*}, Adil Gürsel Karaçor²

Department of Industrial Engineering, Atılım University, Ankara, Turkey¹

Grodan BV, Roermond, the Netherlands²

erman.erk@atilim.edu.tr

Abstract: Large amounts of liquidity flow into several precious metals every day. Investment decisions are mainly based on predicting the future movements of the instrument(s) in question. However, high frequency financial data are somewhat hard to model or predict as stochastic processes and many other random factors are involved. It would be valuable information for the investor if he or she knew which precious metals were quantitatively more predictable, that would also be a good basis for more robust trading decisions. The objective in this study is to build predictive models on high frequency precious metal data and compare predictabilities of different metals using only past price and volume values which should be a basis for robust trading decisions. The data used consist of various frequencies from 1-minute to 4-hour covering a period of almost 20 years for each instrument and frequency. Artificial Neural Network (ANN) and Gradient Boosted Decision Tree (XGB) methods are applied. Comparable results are achieved.

Keywords: PRECIOUS METALS, PREDICTABILITY, ARTIFICIAL NEURAL NETWORKS, XGBOOST

1. Introduction

Large amounts of liquidity flow into the markets all around the world for the trade of various financial instruments such as stocks, commodities, precious metals, foreign exchange (FOREX), futures, and so on every day. Per experts and professionals, average daily turnover in precious markets alone is more than 50 billion US dollars. Investment and trading decisions, whether they tend to be long term or short term, are mainly based on predicting the future movements of the precious metal(s) in question. There has been an ongoing debate among researchers on whether financial markets are predictable or not for a long time. Some think that financial market movements are nothing but random walk, and some findings support that claim: for example, VIX futures prices were found to be unpredictable by Konstantinidi & Skiadopoulos [1]. Furthermore, some researchers even claimed that none of the conventional predictive models proposed in the literature on stock prediction seems capable of systematically predicting stock returns in long range of time horizons, and speculators do not earn significant profits in commodity and interest rate futures markets in aggregate [2, 3, 4]. On the other hand, many researchers disagree with this random walk approach. Some of them claim speculators can gain profits on commodity and currency futures [5, 6, 7, 8, 9]. Some others believe that financial instruments, commodities in particular, are predictable, at least to a certain extent [10, 11]. The debate has not been settled yet; however, it is fair to say that financial instruments are hard to model or predict, if not totally unpredictable. Obviously, it would be valuable information for the investor if he or she could know in advance which way an instrument would go. In this paper an investigation is carried out to give the investor a potential advantage by trying to answer the following question: would a certain precious metal price be likely to go up or down in the following hours? Technical (quantitative) analysis i.e. past price values of the instruments is solely used, and fundamental analysis is not in the scope of this study.



Fig. 1 Sample OHLC price graph for platinum (XPTUSD)

2. Data

The data used in our model consisted of 1-minute (1m), 5-minute (5m), 15-minute (15m), 30-minute (30m), 1-hour (1h), and 4-hour (4h) intraday frequencies covering the period between 2000 and 2019. High frequency data were preferred, because in lower frequencies more and more non-technical fundamental factors might start to affect price movements. Three types of precious metals were considered in this study: platinum (XPTUSD), silver (XAGUSD); and gold (XAUUSD). A sample graph of platinum (XPTUSD) is given in Figure 1.

3. Methodology

Twenty-three out of the twenty-four features used for our prediction model are related to past price movements (past high/low values, etc.), and one is related to seasonality. The output was the prediction of the actual result i.e. the maximum price value into the next eight periods (8x1-minute, 8x5-minutes, and so on).

The chosen state-of-the-art Machine Learning (ML) tools are: Extreme Gradient Boosting (XGB) method and Artificial Neural Network (ANN). Usage of ANNs is quite popular for modeling, prediction, and decision making over financial data and ANNs are regarded as an excellent tool for the purpose [12, 13, 14, 15, 16, 17, 18]. On the other hand, usage of XGB method is very rare [19].

XGB is chosen as one of the prediction methods, mainly because it has gained much popularity in the recent years, due to its overwhelming success in data science competitions [20]. Despite its reputation, XGB method has not yet achieved popularity in the literature accordingly.

ANNs are inspired from biological neural networks, and are well known for their high approximation and modelling capabilities [21, 22]. ANNs are used in a wide range of applications from job satisfaction performance evaluation [23], solving capacitated P-median problem [24] to prediction of aircraft accident occurrence [25]. ANNs are trained in order that they learn a set of input-output data that represent usually a very complex or even undefined function. With sufficient number of hidden layers and neurons, they can model any given input-output relationship [26]. All nodes (artificial neurons) are interconnected, thus form a massive parallelism, and each connection has a weight that changes as the ANN is trained, and also each node has an activation function. There are numerous activation functions, ranging from simple linear functions to various nonlinear ones. The nonlinearity of activation functions enables the ANN to learn even the most complex patterns. In this study; following a parameter optimization, an ANN model is constructed with a multi-layered feed forward network, having 2 hidden layers with 55 and 3 neurons, respectively. Hyperbolic tangent is chosen as activation function. The network is trained by back propagation gradient descent algorithm.

4. Training, Testing and Results

The data are randomized to achieve a fair distribution. 70% of the data is chosen to train the models, 20% for testing, and 10% for validation. Python software [27] is used for ANN and XGB modelling.

Both XGB and ANN models were trained until no improvement was brought about in the cross validation set. One of the main issues concerning machine learning is over-learning or over-fitting problem, in which the system memorizes a certain data set rather than learning it, because of too much training. The performance of the system can be excellent on that data set; however, it performs poorly on different data due to the lack of generalization. The opposite of this issue is the under-learning problem. Therefore, sufficient training or hyperparameter tuning is crucial for an optimum model performance. The criterion of training until improvement in the cross validation set stops is a trade off in this study.

Performances on test sets were averaged to obtain NMSE values. NMSE is mean square error divided by variance of desired output. Being a normalized value, it could easily be used for comparing different instruments of different prices and of different frequencies. Since it is an error term, values closer to zero denote better predictability.

Another statistically meaningful variable we used for predictability performance comparison is the correlation coefficient R. R is used to measure how well one variable fits on another, linear regression wise. In our case, these variables were predicted against desired. R value is calculated by the formula below:

$$R = \frac{\sum_i (x_i - \bar{x})(d_i - \bar{d})}{\sqrt{\sum_i (d_i - \bar{d})^2} \sqrt{\sum_i (x_i - \bar{x})^2}} \quad (1)$$

Where N = number of exemplars in the data set, x = model output, and d = desired output. The size of the mean square error (MSE) can be used to determine how well the network output fits the desired output; however, it does not necessarily reflect whether the two sets of data move in the same direction. For instance, by simply scaling the network output, we can change the MSE without changing the directionality of the data. The correlation coefficient R solves this problem. The correlation coefficient between a model output x and a desired output d is defined by formula (1). The correlation coefficient is confined to the range [-1 1]. When R = 1 there is a perfect positive linear correlation between x and d, i.e. they vary accordingly, which means that they vary by the same amount. When R = -1, there is a perfect linear negative correlation between x and d, i.e. they vary in opposite ways (when x increases, d decreases by the same amount). When R = 0 there is no correlation between x and d, i.e. the variables are called uncorrelated. Intermediate values describe partial correlations. All recorded and calculated performance comparison values in terms of R and NMSE for all metals and all frequencies are given in Table 1 and Table 2 which hopefully could give hints about predictability.

Table 1: Detailed predictability performance comparison for the precious metals on lower frequencies

Instrument	4h		1h		30m	
	R	NMSE	R	NMSE	R	NMSE
XPTUSD	0.36	0.91	0.4	0.94	0.42	0.85
XAGUSD	0.41	0.84	0.44	0.81	0.46	0.79
XAUUSD	0.40	0.84	0.45	0.80	0.36	0.87

Table 2: Detailed predictability performance comparison for the precious metals on higher frequencies

Instrument	15m		5m		1m	
	R	NMSE	R	NMSE	R	NMSE
XPTUSD	0.30	0.90	0.39	0.87	0.36	0.87
XAGUSD	0.37	0.86	0.25	0.94	0.27	0.93
XAUUSD	0.36	0.87	0.38	0.85	0.27	0.93

5. Conclusion

As can be seen from Table 1 that silver (XAGUSD) has the highest overall predictability both R and NMSE wise on lower frequencies. We can say that 30m frequency has the best predictability performance again both R and NMSE wise among the lower frequencies. On the other hand, gold (XAUUSD) has the best 1h performance. As per higher frequencies which are shown in Table 2, there is no clear winner metal. Silver and gold perform similarly in higher frequencies, while platinum is considerably better in 1-minute and 5-minute frequencies. Overall metals perform better on lower frequencies. Therefore, from the investor's point of view, trading decisions based on 1-hour or 4-hour data should probably be more profitable, considering commissions and spread margins are usually more disadvantageous in higher frequencies like 5-minute or 1-minute.

Adding more metals such as copper and palladium could be considered as future work. Of course, building a robust trading strategy using predictability as a metric is also in the scope of future work.

6. References

1. Konstantinidi E., Skiadopoulos G. Are VIX futures prices predictable? An empirical investigation. *IJOF* 27 543–560 (2011)
2. Bossaerts P., Hillion P. Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn? *Rev. of Fin. Stud.*, 12(2), 405–428 (1999)
3. Goyal A., Welch I. A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. *Rev. of Fin. Stud.* Vol. 21 Issue 4, pp. 1455–1508 (2008)
4. Hartzmark M. L. Returns to individual traders of futures: Aggregate results. *JPE*, 95, 1292–1306 (1987)
5. Kho B. S. Time-varying Risk Premia, Volatility, and Technical Trading Rule Profits: Evidence from Foreign Currency Futures Markets. *JFE*, 41, 249–290 (1996)
6. Strozzi F., Zaldivar J.M. Non-linear Forecasting in High-frequency Financial Time Series. *Physica A*, 353 (2005) 463–479 (2005)
7. Taylor S. J. Rewards Available to Currency Futures Speculators: Compensation for Risk or Evidence of Inefficient Pricing? *Economic Record*, 68(Supplement), 105–116 (1992)
8. Wang C. Futures Trading Activity and Predictable Foreign Exchange Market Movements. *J. of Bank. & Fin.* 28, 1023–1041 (2004)
9. Yoo J., Maddala G. S. Risk Premia and Price Volatility in Futures Markets. *JFM*, 11, 165–177 (1991)
10. Campbell J. Y., Thompson S. Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average? *Rev. of Fin. Stu.* Vol. 21, Issue 4, pp. 1509–1531 (2008)
11. Zunino L., Tabak B.M., Serinaldi F., Zanin M., Perez D.G., Rosso O.A. Commodity Predictability Analysis with a Permutation Information Theory Approach. *Physica A*, 390 (2011) 876–890 (2010)
12. Cheng W., McClain B.W., Kelly C. Artificial Neural Networks Make Their Mark as a Powerful Tool for Investors. *Rev. of Bus.* 4 – 9 (1997)
13. Dash R., Dash P.K., Bisoi R. A self-adaptive differential harmony search based optimized extreme learning machine for financial time series prediction. *SAEC Vol.* 19 pp 25–42 (2014)
14. Dutta S., Shekhar S. Bond-rating: a Non-conservative Application of Neural Networks. *Proceedings of the IEEE Int. Conf. on NN* 2 pp 443–450 (1988)

15. Lam M. Neural Network Techniques for Financial Performance Prediction: Integrating Fundamental and Technical Analysis. *Decision Support Systems*, 37 567– 581 (2004)
16. Wedding D.K., Cios K.J. Time series forecasting by combining RBF networks, certainty factors, and the Box-Jenkins model. *Neurocomputing* Vol. 10, Issue 2, pp 149-168 (1996)
17. Xi L., Muzhou H., Lee M.H., Li J., Wei D., Hai H., Wu Y. A new constructive neural network method for noise processing and its application on stock market prediction. *App. S. Comp.* Vol. 15, pp 57-66 (2014)
18. Yu L., Lai K.K., Wang S. Multistage RBF neural network ensemble learning for exchange rates forecasting. *Neurocomputing* Vol. 71, Issues 16-18, pp 3295-3302 (2008)
19. Zięba M., Tomczak S. K., Tomczak J. M. Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction, *Exp. Sys. w. Apps.*, Volume 58, pp 93-101 (2016)
20. Chen, T., & Guestrin, C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22Nd ACM SIGKDD* (pp. 785-794). ACM (2016)
21. Karacor A. G., Sivri N., Ucan O.N., Maximum Stream Temperature Estimation of Degirmendere River Using Artificial Neural Network, *JSIR*, Vol. 66 No. 5 pp. 363-366 (2007)
22. Karacor G., Denizhan Y. Advantages of Hierarchical Organisation in Neural Networks, *IJCAS* Vol. 16 pp 48-60 (2004)
23. Aktepe A., Ersoz S. A Quantitative Performance Evaluation Model Based on a Job Satisfaction - Performance Matrix and Application in a Manufacturing Company. *IJIE*, 19(6), 264 – 277 (2012)
24. Shamsipoor H., Sandidzadeh M.A., Yaghini M. Solving Capacitated P - Median Problem by a New Structure of Neural Network. *IJIE*, 19(8), 305 – 319 (2012)
25. Yeoum S.J., Lee Y.H. A Study on Prediction Modeling of Korea Millitary Aircraft Accident Occurrence. *IJIE*, 20(9-10), 562 – 573 (2013)
26. Hornik K., Stinchcombe M., White H. Multilayer Feedforward Networks are Universal Approximators, *NN*, Vol. 2, pp. 359-366 (1989)
27. Rossum G. Python tutorial, Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI) (1995).

Implementation of UAV for environment monitoring of a Smart City with an airspace regulation by AIXM-format data streaming

Vladyslav Alieksiev^{1,2}, Bohdan Markovych^{1,3}

¹ Lviv Polytechnic National University, Lviv, Ukraine

² vladyslav.i.alieksiev@lpnu.ua,  <https://orcid.org/0000-0003-0712-0120>

³ bohdan.m.markovych@lpnu.ua,  <https://orcid.org/0000-0002-8813-9108>

Abstract: One of the most expected and the most desired implementation of modern technologies is the concept of a Smart City. Meanwhile there are many buzz-words known to describe the ability of implementation of many approaches to provide control for Smart City. Many of such technologies had become known for Industry 4.0 concept, including 5G mobile Internet and autonomous aerial vehicles (so called drones). Nevertheless, there is an open question of how to provide and support limits where and when UAV (unmanned aerial vehicle) should be able or not to fly over the city. Naturally, the self-driven drones equipped with a batch of sensors and able to avoid some obstacles would not be enough for everyday use within Smart City. Current paper is focused on discussion of possible solution to provide UAV with up-to-date data on restrictions, that may be required in a modern city.

Keywords: SMART CITY, ENVIRONMENT MONITORING, AIRSPACE REGULATION, INDUSTRY 4.0, UAV, AIXM

1. Introduction

A modern understanding of a 'Smart City' had evolved actively during recent years and implies implementation of most contemporary integrated information and communication technologies to produce, develop, improve and yield common wealth to a local community. Coming initially (according to [1]) from idea of sustainability to cities via providing its development with leading technological solutions, the concept had become later a natural part within a wider concept of a Fourth Industrial Revolution [2] or Industry 4.0.

Nowadays, it is actually impossible to see a strict border between modern city and modern manufacturing in the same way like it was possible even in previous 20th century. The 'Smart City' should be now considered as a big quasi-industrial area which produces, supplies, supports and consumes huge amounts of products and services. The problem of convenient lifestyle had gone outbound of simple concept of dwelling, and it becomes more and more complex in each next decade. Technologies, and ICT in particular, permeate deeper and deeper into everyday life of each citizen and surround the communities in a new transparent manner, giving a new sense of collaborative existence for the whole city. One of a key modern aspects and a technology that changes the face of any our city into the 'Smart City' is named 'Cyber-physical System' (CPS). CPSs bring smart automation into our life. A significant role in integration of CPSs into Smart City will take an army of autonomous flying robots (in different sources there were different terms and abbreviations used, like UAV for 'Unmanned Aerial Vehicles', or UFV for 'Unmanned Flying Vehicles', or UAS for 'Unmanned Aerial System'), known also simply as 'drones'. These smart machines should serve people and cities for many known purposes, but in a new way.

There is a number of already known and actively developed smart robots which came to modern cities already. Among such smart solutions there are delivery robots (see Fig. 1) – autonomous wheel robots from Estonian company "Starship Technologies" [3]: "According to the concept, those robots will deliver everything from shopping bags to internet purchases, pizzas and many other things, freeing us from wasting time and energy on such daily chores". Surely, these kind of robots are pioneers in their field and developers know themselves of their main problem [3]: "As the robot will just be delivering the package to the front of the address and not bring it up in an elevator or staircase, it would be pretty strange to order a pizza in a Hong Kong high rise and to then discover that the whole front of the building is full of robots".

The more promising, from perspective of mobility, among autonomous robots are drones, the flying brothers of a street wheel robots. Fortunately, we also have a number of examples of successful implementations in the area already.



Fig. 1 Delivery robot from Estonian company "Starship Technologies" in urban landscape (image retrieved from [3])

The concept of air deliveries was actively developed in recent years by Amazon [4, 5]. The first fully autonomous air drone delivery from Amazon Prime Air was announced to be made on December 7, 2016 [4] (see Fig. 2).



Fig. 2 Delivery drone from Amazon Prime Air (image retrieved from [4])

And Amazon didn't stop with that, presenting their new fully electric delivery drones in June, 2019 [5] (see Fig. 3). These drones looked different to the previous generation of delivery drones. So, the fleet of delivery drones includes quite different flying machines able to carry different kind of parcels of different sizes and weights.



Fig. 3 Presentation by Jeff Wilke of fully electric delivery drone from Amazon Prime Air at “re:Mars” conference in Las Vegas (image retrieved from [5], photo by Jordan Stead, Amazon)

However, UAV and wheel robots for delivery services can be the first ones in the airspace and in the streets of Smart Cities, but not the only ones to serve people with their needs. Let’s have look at industrial implementations of UAV, and purposes and services provided with such drones. In December, 2016 Audi had presented their vision for the future of their car manufacturing and one of their innovations was to involve flying drones to transport parts to the particular place for the conveyor [6, 7] (see Fig. 4).



Fig. 4 Audi smart factories using drones to transport parts (image retrieved from [6])

A great benefit from flying drones was found for smart agriculture. India has a successful experience in that area: “The aim of the Syngenta-Skylark project was simple: count the number of germinated seedlings and plants across corn farms using Syngenta seeds with an intent to see how to increase productivity” [8]. Another successful drone usage for monitoring in agriculture and implementation of precision agriculture according to [8]: “... company Cyient and a leading fertilizer company in south India are working together on a precision agriculture project to monitor farms using drones in order to monitor and act on crop health. ... The idea is to identify areas that are stressed, either due to a nutrient deficiency or affected by weeds or pests, and alert farmers in the form of digital prescriptions.” All these implementations revealed a significant impact on quality of agricultural processes not only in way of looking at the situation in a whole and making some decision, but also in a way of supplying exact quantities of chemicals to particular plants and not to make unnecessary influence and chemical pollution for the rest of normal growing plants (see Fig. 5).



Fig. 5 Drones used for agriculture in India (image retrieved from [8])

One of the world’s biggest steelmakers declared the use of drones to provide more safe industrial manufacturing: “New ways of incorporating drones into the workforce are popping up everywhere, such as for marketing, expedient delivery and information gathering. In heavy industries like steel manufacturing, drones are playing an increasingly important role in workplace safety. When paired with AI, drones have even more possibilities for minimizing workplace hazards.” [9] (see Fig. 6). And also more precise, agile and safe monitoring is available: “...steelmaker POSCO is incorporating drones into its smart factory to monitor gas leaks that are hard to detect” [9].



Fig. 6 Drones used for monitoring at steel manufacturing company POSCO, South Korea (image retrieved from [9])

Drones are used in factory automation for critical safety issues: “The interconnected and interdependent nature of the bulk power system requires a consistent and systematic application of risk mitigation across the entire grid system to be truly effective” [10] (see Fig. 7).



Fig. 7 Drones used for monitoring of industrial safety in USA (image retrieved from [10])

Current drones' areas of application are widely known: "UASs of all types have already been used in a wide variety of applications in practical ways, such as aerial photography, agriculture, commercial delivery, entertainment, exploration, national defense, public safety, surveying, and thermography" [10].

Obviously, growing sets of features and abilities, that drones can provide to different industries, are very promising for development in number of new areas, like Smart City needs. This growth leads also to inevitable question of coordination and regulation of drones' flights. The focus of our current research is the problem of necessity of regulation of UAV flights within modern smart city.

2. Prerequisites and means for solving the problem

Modern city can be much more complex if to compare with above mentioned use cases of industrial implementations of drones. Besides deliveries in an urban area, among the possible uses for UAV in Smart City there various monitoring purposes that not just imply but require all of the following and even more:

- People support and crowds monitoring – similar to personal workers' safety support in steel manufacturing.
- Safety and accident monitoring – same as in steel manufacturing (or any other dangerous manufacturing processes).
- Leaks monitoring – same as in industrial dangerous gases and liquids supply systems.
- Power supply lines and power grids monitoring – exactly the same as in power grids systems.
- Plants growing and monitoring – same as in precision agriculture.
- Animals monitoring – same as in precision agriculture and breeding.

Considering these and many other possible implementations of UAV useful abilities, one can conclude that at least two important aspects would be brought with UAV utilization in Smart City:

1. Changes in citizens' lifestyle.
2. Management for UAV as a whole system.

Example for the first aspect can be found in area of delivery robots and drones. Currently, most citizens are visiting supermarkets, big stores or small shops to buy goods, food etc. The city transport infrastructures are used to provide people with ability to go between their homes and places where they buy anything. Even in case of making orders by phone or in Internet, many delivery services use courier services to deliver purchases to clients and again city transport infrastructure suffer with additional load (due to growing demand). During rush hours these deliveries will also yield the growth of delays. These delays impact not only delivery services but also public transport and personal vehicles. Now, providing citizens with efficient delivery services utilizing UAV, the impact on on-ground infrastructure and the lifestyle will change: citizens will reduce their need to spend time in traffic jams and will gain a little bit more of free time for other activities.

The second aspect comes from creation of new infrastructure (the infrastructure for UAV) and from utilization of new spaces (the airspace) and facilities (take-off and landing, fueling or charging, maintenance and support). Meanwhile, UAV should avoid collisions and excessive airspace congestion (due to inevitable unwanted crashes and noise pollution). Another big problem is an impact on city ecosystem (birds and many other flying species will suffer more and more from human activity). Evidently, all this complexity leads to demand of management and regulation for the whole new system of UAV within Smart City.

Our current suggestion is to develop the idea and approach for industrial implementation of airspace regulation [11] for the Smart Cities and, in particular, for the purpose of environment monitoring. The experience of aviation in data exchange, like AIXM [12], and

rather difficult cases of governance in airspaces performed by international organizations, like EuroControl [13], can provide Smart Cities with a good examples of airspace management. In example, if compare UAV, which are self-driven, to aircraft flights we suggest to use VFR and IFR flights as analogues for drones to fly relying on own sensors only or to fly following some given prescriptions. Hence, based on known experience, some problems could be identified and solved initially.

One of the key problems here is the problem of data exchange to provide UAV with up-to-date restrictions. Unlike to aviation, there is no pilot to speak to dispatchers, who can provide flight support from the ground. Also, the AIXM data format is based on XML notation supplemented with GML extension [14], and the source files volume requires huge channel capacity to perform data transmission. While each millisecond in flight is precious, data transmitting time measured in seconds or minutes looks feasible only before take-off and after landing, but not during the flight. As an example, the set of source files describing navigation data and flight conditions for Europe has volume about 1 Gb in original text format (in XML notation) or about 80 Mb in compressed format (zip-packed files). Considering the number of objects and possible limitations in modern city, the description of a single city airspace wouldn't have less volume of data.

3. Solution of the examined problem

As one can calculate easily, deployment of 1 Gb of data arises many problems:

1. Is it worth uploading the full data set for the flight before take-off? This is appropriate for short time flights, but what about a long lasting ones?
2. How can city provide UAV with immediate updates of data during the flight?
3. Would it be efficient to rely on external data storage and allow network access (via 4G/5G wireless network)?

It looks obvious, that single solution cannot provide the required reliability and a combination of approaches should be used.

Let's consider the problem of data volume and an approach to overcome this obstacle. Uploading data into UAV memory (data storage) requires regular upgrade of memory blocks to increase the capacity. On one hand, this is not a big problem, while UAV should have maintenance and support regularly. On the other hand, it reduces the ability of in-flight automatic regulation. For example, when some accident happens and an airspace should be used by emergency services drones only and not by delivery drones, how should the city dynamically provide all the drones with the restriction? This requires an update, which can be made via wireless network (WiFi hot spots or mobile 4G/5G) from a data server.

The opposite case of relying on data transmissions only looks also inappropriate due to possible communication problems via wireless networks etc. So, the UAV will require to upload most current dump of data before take-off and during the flight to check for updates (itself) or being securely forced to update data (via some kind of notifications). Here again the UAV faces the problem of data volume and time losses. Downloading data in a raw XML format will require a time consuming parsing to "understand" data. Also, if to download a complete data set it will produce a rather big time delay to deploy at UAV. But if to download a portion of data, it may also take some time to parse it and to find the right place to deploy it (replace existing data fragment or add new one). Here the good known technique of remote database can help. Our suggestion is to rely in this case on relational databases, which are a good developed tool for operations with data, which can be updated often (unlike to no-SQL databases, which showed good in often data selects but are less useful in situation with often updates).

To solve this problem, we had developed a software to convert plain text of XML based AIXM format into relational database.

Deploying such database on both UAV and data server gives a good opportunity to perform data exchange in more compact format, than the original AIXM format provides. Nevertheless, the database schema follows the detailed structure of AIXM format and allows to keep all the variety, that AIXM brings. Evidently, that utilization of SQL queries makes easier both development of software for UAV and data servers of Smart City.

Using this technique allow to build and develop the structure of urban airspace. Flying drones are able to provide an up-to-date information about current situation in the city and make it possible to perform dynamic regulation of city airspace. Also, an automated invoke of drones' assistance and request for services becomes available. The optimized for relational database AIXM data format allows to set up efficient data exchange and streaming of updates in a quick and easy to use way.

4. Results and discussion

Our experiments were made with navigational data published for aviation purposes by EuroControl and provided by RocketRoute company [15]. The data conversion takes two major steps:

- Parsing XML and preparing SQL statements to create data structure and insert data – took about 114 seconds on approximately 1 Gb of source files (data volume in SQL format reduced to approximately 800 Mb – about 20% in compare to original AIXM notation).
- Executing SQL statements to deploy the database – took about 354 seconds (the volume of database with indexes raised to approximately 1.5 Gb).

The implementation of described ideas and software applications requires further research with city circumstances and particular drone fleet.

5. Conclusion

We offer to implement known industrial standard from aviation to provide Smart City with ability to manage airspace and establish unified control for UAV.

Using the techniques discussed above will allow to establish a reliable urban environment monitoring systems and provide many different applications of UAV to support industrial services within modern cities.

6. Acknowledgement

This research was sponsored by RocketRoute Limited (London, United Kingdom), including current problem analysis and software development. Authors give thanks to Uwe Nitsche, CEO of RocketRoute Ltd. (London, UK), for supporting this research by granting access to some company's databases (airspace network, navigation data, etc.) and making the case study possible. The developed software is deployed on servers and remains the intellectual property of RocketRoute Ltd. (London, UK).

7. Literature

1. Smart Cities: Contradicting Definitions and Unclear Measures / M. Cavada, D. Hunt, C. Rogers // 4th World Sustainability Forum 2014: Conference Proceedings Paper, November 1–30, 2014. – 12 p. – <http://sciforum.net/conference/ws-f-4/paper/2454>.
2. K. Schwab. The Fourth Industrial Revolution. – NY: Crown Publishing Group, 2016 (2017). – 192 p.
3. Toivo Tänavsuu. Estonian delivery robots are transforming the world // Estonia.ee – January 2017. – <https://estonia.ee/delivery-robots-created-by-estonian-engineers-are-transforming-the-world/>
4. Prime Air // Amazon.com – December 7, 2016. – <https://www.amazon.com/Amazon-Prime-Air/b?ie=UTF8&node=8037720011>
5. Frederic Lardinois. A first look at Amazon's new delivery drone // TechCrunch.com – June 5, 2019 – <https://techcrunch.com/2019/06/05/a-first-look-at-amazons-new-delivery-drone/>
6. Jamie Harris. From 3D-printed metal to drones that transport parts: this is how your future car will be built in smart factories // BT.com – December 3, 2016 – <http://home.bt.com/tech-gadgets/future-tech/audi-smart-factory-drones-3d-printing-11364114919294>.
7. Audi Smart Factory – Future of Audi Production // YouTube, 17.11.2016 – <https://youtu.be/sqCbYd8O8MU>.
8. Anand Murali. How drones are paving a new path for precision agriculture in India" // Factor Daily, 04.07.2018 – <https://factordaily.com/drones-for-precision-agriculture-in-india/>.
9. Why More Drones Will Fly the Skies of Workplaces" // POSCO Newsroom, 15.12.2017 – <https://newsroom.posco.com/en/drones-will-fly-skies-workplaces/>.
10. DRONES in automation / Marissa Morales-Rodriguez, Sterling Rooke, Peter Fuhr, and Penny Chen, // ISA, InTech Magazine, 2017 – <https://www.isa.org/intech/20170603/>.
11. Alieksieiev V. An approach of FRA and its use for UFV supported manufacturing / V. Alieksieiev, B. Liubinsky, V. Hladun // "Machines. Technologies. Materials." – Scientific-Technical Union of Mechanical Engineering "Industry 4.0" (Sofia, BULGARIA), 2019 – Vol.13, Issue 10/2019. – P437–438. – <https://stumejournals.com/journals/mtm/2019/10/437>.
12. Aeronautical Information Exchange Model – <http://aixm.aero>
13. EUROCONTROL – <https://www.eurocontrol.int/>
14. Geography Markup Language (GML) – https://en.wikipedia.org/wiki/Geography_Markup_Language.
15. Flight Planning Centre RocketRoute – <https://fly.rocketroute.com>.

E-university cloud information security issues

Firudin Aghaev¹, Gyulara Mamedova², Esmira Aleskerova³
 AMEA, Institute of Information Technologies, Baku, Azerbaijan
 agayevinfo@gmail.com, gyula.ikt@gmail.com, aleskerova-66@mail.ru

Abstract: This article identifies various security problems of electronic education in the provision of cloud services and proposes solutions to ensure security measures. Various types of attacks on e-learning platforms are also discussed. Various models of the usage of cloud technologies in electronic education, threats and security requirements when using these models are investigated.

Keywords: ELECTRONIC UNIVERSITY, CLOUD SERVICE DELIVERY MODELS, IDENTIFICATION AND AUTHENTICATION CONTROL, CLOUD SERVER PROTECTION, MITM ATTACKS, DDOS ATTACKS, INSIDER ATTACKS.

1. Introduction

Cloud-based e-learning is one of the fastest growing information technologies that offers powerful cloud-based e-learning products [1-2]. Cloud technologies have numerous advantages over existing traditional e-learning systems, but at the same time, security is a serious problem in cloud e-learning. To prevent the loss of valuable user data due to security vulnerabilities, appropriate security measures must be followed. Cloud-based e-learning products must meet customer security needs and address various security threats. In this paper, we consider key security issues when using cloud computing in e-learning systems.

2. E-learning cloud architecture

The cloud architecture of e-learning is mainly divided into five levels [3] called: the level of hardware resources, the level of software resources, the level of resource management, the server level and the application level. The cloud architecture of e-learning is illustrated in Fig. 1.

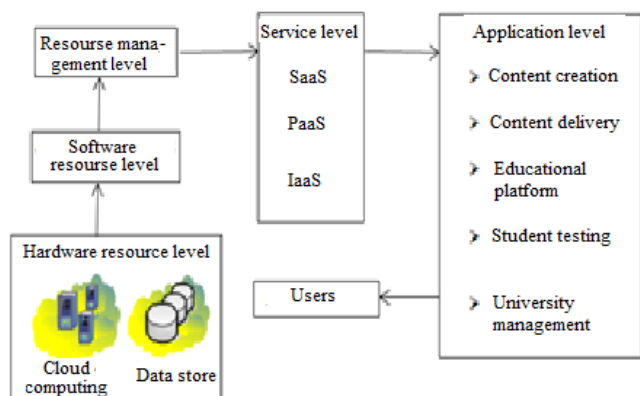


Fig. 1. Cloud architecture of e-learning.

The level of hardware resources is the lowest level of cloud infrastructure, which provides the computing power of electronic education. To ensure the smooth operation of cloud services for e-learning systems, physical memory is dynamically expanding, and is scaled at any time to add additional memory.

The software resource level includes an operating system and application software for developers and users of the cloud infrastructure.

The level of resource management plays an important role in the weak relationship of software and hardware resources. Using the idea of virtualization, it provides uninterrupted distribution of software on demand for various hardware resources of the cloud infrastructure.

The service level is divided into three levels: IAAS (Infrastructure as a Service), PAAS (Platform as a Service) and SAAS (Software as a Service). These service levels help cloud clients use various types of cloud services, such as software, hardware, and infrastructure services.

The level of applications differs from all other levels in the cloud-based e-learning architecture, as it forms the basis of components for e-learning and serves to create training content, content delivery, educational platform, assessment of learning and education management.

3. Problems of ensuring services of cloud calculations for electronic education

As was shown above, there are three models of cloud service delivery: 1. Software as a service (SaaS), 2. Platform as a service (PaaS), 3. Infrastructure as a service (IaaS). This article addresses security issues in each of these groups.

Software as a Service (SaaS) - allows multiple users to work together with ready-made software. Therefore, in SaaS, the user is completely dependent on the software vendor. To ensure proper security measures, the software vendor must ensure the security of the training data so that it is not possible for several users to see each other's data. The educational institution must be sure that the cloud service provider will only process data in accordance with its instructions, that it will take appropriate measures to exclude unauthorized access to the data, its modification or destruction.

Platform as a Service (PaaS) provides a computing platform and system software as a service. PaaS provides services for software developers. These include platforms such as Google App Engine, Salesforce, Windows Azure and others, which allow you to create programs and corporate sites in Java and Python. PaaS providers provide services for application development, deployment, team collaboration, web services integration and testing [4].

The main security threats to the PaaS layer are data location and privileged access. In the United States and many EU countries, universal safety standards and data privacy laws have been set for data placement problems. For example, the General Data Protection Regulation (GDPR) of April 27, 2016, valid in all EU countries, they never allow sensitive data to move from the country [5-6]. For violation of the rules for processing personal data, fines reach millions of euros.

Based on the location of the data, the PaaS model provides reliability for its customers. To ensure reliable data storage at the PaaS level, the user must choose a reliable encryption method for accessing data, maintain high standard data confidentiality, require legally executed contractual obligations of security mechanisms from the cloud provider. The cloud service provider must have technical solutions to prevent unauthorized access for users and support the principle of separation of duties for privileged users in order to prevent and detect malicious insider activity.

When storing encrypted data in a cloud storage, decryption keys must be stored securely in other disintegrated systems [7-8]. If the cloud user system allows the cloud service provider to process unencrypted data, the cloud service provider must ensure that the data is protected from unauthorized access, both inside and outside.

Infrastructure as a Service (IaaS) allows you to use many resources, such as servers, storage, networks, and other computing resources; it is hosting for a virtual machine. In the event that IaaS provides complete control and management of resources, users can

safely run any software on dedicated resources. In this case, an agreement is reached between the user and the cloud provider on the SLA (Service Level Agreement), which describes the services provided, the rights and obligations of the parties. This agreement contains a detailed description of the services, methods and controls provided to ensure the security of information.

The main types of threats in electronic education when providing IaaS services are DDoS attacks (Distributed Deny of Service Attack), MITM attack (Man In The Middle), DNS attacks, etc. The purpose of a DDoS attack is to make a cloud service unavailable for an auto-customized user. In this case, the cybercriminals use SYN flood (TCP connection requests), in which the entire server channel is simply clogged with connection requests. SYN flood is one of the types of denial of service network attacks, which consists in sending a large number of SYN requests, overflowing the connection queue on the cloud server.

Another common threat while providing IaaS services is the MITM attack, when an attacker introduces himself between two legitimate users of the network. An attacker establishes a connection between two users and tries to take possession of the information sent by them to each other [9].

Recently, in a cloud environment using the IaaS service, hackers have used this kind of

instantly change the parameters of the transaction, as well as the pages of the user's request, is completely transparent to the victim.

4. Safety measures of cloud computing services for electronic education

While providing cloud services for an electronic university, to ensure the security of information, special attention needs to be paid to protecting hardware and virtual data processing devices, as well as communication channels. It is necessary to provide the following safety criteria [10-11]:

- Control of identification and authentication of subjects and access objects;
- Protection of computer storage media;
- Ensuring the necessary level of cryptographic and anti-virus protection of stored and transmitted information;
- Protection of the cloud server, communications and data transmission;
- Adoption of firewall measures.

In order to protect data from confidential information leakage, it is necessary to use strong encryption methods for network traffic to control data flow in the network: Secure Socket Layer (SSL) and Transport Layer Security (TLS). These security levels will help protect against traditional network problems such as MITM attacks, IP spoofing, port scans, sniffing packages, etc.

Attacks today have become complex and multi-level, evil-wills choose specific targets and take a long time to prepare for an attack in order to hit precisely the most vulnerable elements. The evolution of Internet threats has naturally influenced the development of a means of countering them. Today, the most effective tools for protecting cloud resources from external attacks from the network are Cloud Access Security Broker (CASB) solutions aimed at satisfying new security requirements. CASB is a unified tool for monitoring all cloud applications, resources and services, controls the interaction between cloud applications (access, traffic, downloading and data storage), the server and external users, allows you to identify potential threats and is focused on a high level of protection cloud environment [12-13]. CASB prevents unauthorized actions of users, detects abnormal activity, including those associated with the actions of various malicious programs. According to data published by Gartner in 2019, the leaders in the

cloud security market are: Skyhigh Networks, Netskope, Symantec [14-15].

And the recently appeared CloudSOC Security for Cloud Apps software package [16] monitors in real-time transactions with authorized (permitted) and unauthorized cloud applications; implements a visualization of user activity maps for quick analysis of their actions; protects against threats based on extensive analytics of user behavior. It provides rich visibility, data movement control and sophisticated analytics to identify and combat cyber threats in all of your cloud services.

The use of CASB tools in electronic education will make it possible, by analyzing the logs of "registration" and "geolocation", data on the time of performing certain actions in the "cloud", to track user actions, identify threats in real time, and identify unsafe applications. Machine learning tools are built into these packages, which provide the ability to track user behavioral actions and their deviation from the norm. By analyzing risk indicators, such as: insecure IP addresses, registration failures, activity level, inactive accounts, "impossible travel" scenarios and their location, timely detection of intruders on the network.

5. Conclusion

The problem of data accessibility in electronic education is the main obstacle to ensuring the security of cloud data. A literary study clearly shows the risks in e-learning based on cloud computing, as well as its service delivery models and effective solutions for each attack. The listed security issues are important for management and the new methodology for developing secure e-learning based on cloud computing in the future.

The main goal of the work is to elucidate the key security problems that arise when implementing cloud computing for e-learning systems. The development of e-learning systems should be carried out using safety methods and internationally recognized standards. The system must implement security services, such as authentication, encryption, access control, user management and their permissions.

6. References

- [1] N. Antonopoulos, L. Gillam. Cloud Computing: Principles, Systems and Applications. London: Springer-Verlag, 2010. 379 p
- [2] H Skleyter. Oblachniye vichisleniya v obrazovanii. Analiticheskaya zapiska/perevod s angliyskogo. Institut YUNESKO po informazionnim texnologiyam v obrazovanii [Elektronniy resurs] // IITO YUNESKO: 2010. URL: <http://iite.unesco.org/pics/publications/ru/files/3214674.pdf>.
- [3] K.Fogarti. Oblachniye vichisleniya: opredeleniya i resheniya. [Elektronniy resurs] // Izdatelstvo Otkritiye sistemi. <http://www.osp.ru/> URL: <http://www.osp.ru/cio/2011/03/13007508>.
- [4] V.F. Shangin. Informaionnaya bezopasnost kompyuternix system I setey / V.F. Shangin. – M.: IZD «FORUS»: Ivfra-M, 2008. – 416 s.
- [5] https://ec.europa.eu/info/law/law-topic/data-protection_en
- [6] A. Pazyuk, M. Sokolova. Zashita personalnix dannix: mejdunarodniye prinzipi i standarti. <https://www.researchgate.net/publication/281459857>
- [7] A.V. Erigin. Analiz effektivnosti sistem predotvrasheniya utechek konfidenzialnoy informazii iz lokalnix setey. Vestnik SibADI, vipusk 2 (20), 2011, s. 40-47.
- [8] M.Jensen, J.Schwenk, N.Gruschka, & L.Iacono, Year. On technical security issues in cloud computing. In, 2009. IEEE, pp.109-116.
- [9] Callegati, Franco; Cerroni, Walter; Ramilli, Marco. IEEE Xplore - Man-in-the-Middle Attack to the HTTPS Protocol (англ.) // ieeexplore.ieee.org : journal. — 2009. — P. 78—81.
- [10] E. B. Fernandez, Nobukazu Yoshioka, and Hironori Washizaki, "Patterns for cloud firewalls", Procs. of AsianPloP (AsianPattern Languages of Programs) 2014, Tokyo, Japan, March 2014
- [11] E.A .Isayev., D.V. Dumskiy., V.A. Samodurov., V.V.Kornilov. Obespecheniye informazionnoy bezopasnosti oblachnix sistem Matematicheskaya biologiya i bioinformatika , 2015. T. 10. № 2. s. 567–579.

- [12] Hassan Reza, Madhuri Sonawane, "Enhancing Mobile Cloud Computing Security Using Steganography", Journal of Information Security, July 2016, Page no. 245-259.
- [13] https://www.anti-malware.ru/analytics/Market_Analysis/cloud-access-se-curity-broker
- [14] <https://www.gartner.com/reviews/market/cloud-access-security-brokers/co-mpare/symantec-blue-coat-vs-skyhigh-networks>
- [15] C. Lawson, S.Riley Magic Quadrant for Cloud Access Security Brokers. Gartner. 29 October 2018.
<https://www.bsigroup.com/globalassets/local-files/enie/csir/resources/whitepaper/1810-magic-quadrant-for-casb.pdf>
- [16] <https://www.symantec.com/content/dam/symantec/docs/data-sheets/cloud-soc-gateway-en.pdf>

Personal data security techniques in e-education

Firudin Aghaev¹, Gyulara Mamedova², Rena Melikova³
AMEA, Institute of Information Technologies, Baku, Azerbaijan
gyula.ikt@gmail.com, rena22@rambler.ru, lalamailala@bk.ru

Abstract: In recent years, the popularity of electronic education is gaining momentum. In this regard, there is a problem of ensuring its safety, which should be carried out using safety methods and internationally recognized standards. This article identifies various problems of personal data security in electronic education and suggests solutions to ensure the protection of educational information.

Keywords - ELECTRONIC UNIVERSITY, PROTECTION OF EDUCATIONAL INFORMATION, CROSS-SITE SCRIPTING, SQL ATTACK, SECURE AUTHENTICATION, DIGITAL SIGNATURE

1. Introduction

The institution uses a variety of information that requires protection. This is the personal data of students, teachers, administration and other categories of users. This also includes information constituting a commercial secret of the university (educational programs and study materials, results of research work, etc.), which allows it to stay ahead of other universities in the field of providing better education, more advanced teaching methods, and better educational programs.

Research and consulting company Gartner [1], specializing in information technology markets, predicts that the cost of information security in the world will exceed \$ 124 billion in 2019 and will affect various segments, such as identity and access management (IAM), identity governance and administration (IGA) and data loss prevention (DLP).

Compliance with security requirements in the e-learning system is an extremely difficult problem, because it is necessary to protect content, services and personal data not only for external users, but also for internal users, including system administrators. The training system should implement security services such as authentication, encryption, access control, data integrity, content protection and user management. In this article, we will look at some key security issues that need to be considered when designing and using e-learning.

2. Electronic university information security system

The main methods for ensuring the security of information in electronic education are [2-4]:

- Organizational remedies. Used to restrict access (excludes access to information of unauthorized persons); access control - the separation of information into parts and the organization of access to it in accordance with the functional responsibilities and authority of the user; access control - determining the authenticity of a subject who has access to information.

- Hardware protection. These include protection against server malfunctions, protection against malfunctions of information storage devices, protection against information leaks, electromagnetic radiation.

- Software protection - to identify technical devices and programs that pose a danger to the normal functioning of electronic education.

A secure training platform should include the following key aspects of security:

- the availability of training information (the ability to obtain the required information in a reasonable amount of time);
- integrity (protection of educational content from destruction and unauthorized changes);
- confidentiality (protection of personal data, educational content and system management from unauthorized access).

The main elements of the implementation of threats to educational data in the computer network of the university can be

objects or entities that create these threats, the communication environment for the distribution of personal data, media (a material object in which personal data is reflected in the form of text, video and sound information).

Data security is closely related to the integrity of programs and operating systems. If the integrity of the operating system is violated, then the control monitor may stop working properly. A control monitor is a mechanism that ensures that only authorized entities can access data and perform operations. Obviously, information security cannot be guaranteed if the mechanism for checking and restricting access to data does not work. For this reason, in order to protect the data itself, it is important to protect the integrity of the operating systems. [5].

Secure authentication is required to identify the user and determine their access rights when using web applications. This mechanism does not allow attackers to gain access to another user's account, view confidential information or perform unauthorized operations. In addition, after authentication, the user should be able to change his password

Access control during authentication restricts access to the system of unauthorized users and allows the user to perform only their allowed operations in the system (administrator, editor, instructor, student, registered user, unregistered user, etc.).

3. Methods of protection of educational information in electronic education

When introducing an e-learning system in an educational institution, it should be checked for external intrusion problems, such as:

- Cross-Site Scripting XSS;
- Remote injection using a virus / trojan file;
- SQL injection to the site address (URL SQL injection);
- Hacking passwords using decryption systems;
- Guessing the identifier of the website session (session forecast).

Cross-Site scripting is one of the most common application level web attacks. XSS is usually intended for scripts embedded in the page, which are executed on the client side (in the user's web browser) [6-8]. XSS itself is a threat caused by the weakness of the Internet security of scripting languages. The concept of XSS is to manipulate client-side scripts on a web application so that they run as an attacker wants them to.

Such manipulation can embed a script in a page, which can then be executed each time the page is loaded, or whenever the corresponding event is executed. An XSS attack can be used to achieve the following objectives:

- access to confidential information;
- identity theft;
- change browser functionality;

- damage to web applications;
- denial of service.

Figure 1. shows the most vulnerable areas of the spread of threats to educational data in the computer network of the university [9].

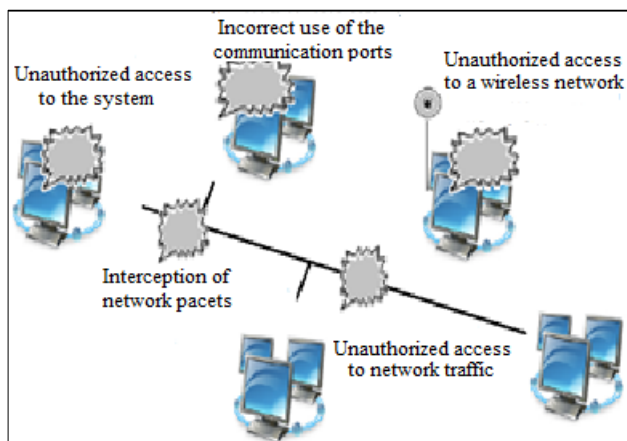


Fig. 1. The most vulnerable places in the computer network of the university for the spread of threats to educational information.

Denial of service means that the user (student or teacher) cannot perform the corresponding actions in the system. Suppose a teacher deletes his student's exam results. In this case, it should be possible to track who deleted them using some script file. These script files must be reliable and tamper-proof. An audit mechanism is used to fulfill this requirement.

To prevent such an attack, it is necessary that the developer of the e-learning platform ensures the safety of using the web page so that the pages on the website return user data only after checking for malicious code. It is necessary to make extensive use of testing tools at the design stage in order to eliminate the possibility of XSS attacks in e-learning applications before they are put into operation.

The next type of attack is SQL attack [10-12]. Using this attack method, hackers inject SQL queries or characters into a web application in order to gain unauthorized access to the database. Such requests may result in access to unauthorized data, bypassing authentication. This threat can be avoided by strictly following some basic coding practices. The most common methods to prevent this vulnerability are:

- checking when entering user SQL queries for dangerous characters, for example, single quotes;
- encryption of confidential data;
- ensuring that error messages do not notify unwanted users about the internal architecture of the application or database.

An SQL attack can also be applied to URLs [13], which can be modified by an attacker to access sensitive information. To prevent this, it is advisable to avoid sending important parameters to the URL. This is achieved by transmitting a unique and hard to guess value of the identifier (session identifier), which the browser sends with each new request either to the cookie or to the URL. Sessions are a way to store state and user variables for subsequent page requests. The session is alive as long as the browser continues to send an identifier with each new request. Session prediction means guessing a valid session identifier using various tools and methods (e.g. brute force technique). An attack is possible when the session identifier is weakly encrypted, too short, or assigned sequentially.

Sessions that do not expire on the HTTP server can allow an attacker to guess or try out a valid authenticated session identifier for an unlimited amount of time and ultimately gain access to that user's web accounts. In addition, the session identifier can be

registered and cached on proxy servers. When transmitted via URLs, GET requests can be stored in browser history, cache and bookmarks that can be seen. To prevent session security issues, you should follow these guidelines:

- The session identifier should be long enough and unpredictable;
- verify the correctness of the session identifier;
- avoid the option "remember me" (permanent logins);
- stop the session when a security error is detected;
- stop the session after a period of inactivity;
- delete cookie session when the session is destroyed.

Cross-site replication, which also improves data processing speed, is a good practice for ensuring data security and integrity [14].

4. Conclusion

In this article, we described some aspects of the security of e-learning platforms and, in particular, analyzed the most important problems, such as Cross-site Scripting, SQL attacks, password cracking, etc. The development of e-learning systems should be carried out using safety methods and internationally recognized standards. The system must implement security services such as authentication, encryption, access control, manage users and their permissions. Data transfer between the system and administrators or content operators must be done via encrypted SSL channels through the web administration interface.

5. References

- [1] <https://www.gartner.com/en/newsroom/press-releases/2018-08-15-gartner-forecasts-worldwide-information-security-spending-to-exceed-124-billion-in-2019>.
- [2] V.V. Gafner, Informazionnaya bezopasnost / V.V. Gafner. - M.: Feniks, 2017. 336 s.
- [3] P.P.Urbanovich. Informazionnaya bezopasnost i nadejnost sistem - Minsk: BGTU, 2007. 90 s.
- [4] A.V.Vasilkov. Bezopasnost I upravleniye dostupom v informazionnix sistemax / A.V.Vasilkov, I.A. Vasilkov. - M.: Forum, 2017. 368 s.
- [5] G. Gagne, P. B. Galvin, A. Silberschatz? Operating System Concepts, Publisher: John Wiley & Sons, 2012, p. 312.
- [6] Z.A.Nosirov, I.M.Ajmuramedov. Obnaruzheniye XSS-uyazvimostey na osnove analiza polnoy karti web-prilozheniya. Sistemi upravleniya, svyazi I bezopasnosti. №1, 2018, s. 78-92.
- [7] S. Fogie, J. Grossman, R. Hansen, A. Rager, P.Petkov. XSS attacks: Cross Site Scripting exploits and defense - Seth Fogie, Oxford: Elsevier Limited, 2007. 448 p.
- [8] M. Denis, C. Zena, T. Hayajneh, "Penetration Testing: Concepts, Attack Methods, and Defense Strategies", ISBN: 978-1-4673-8490-2, DOI: 10.1109/LISAT. 2016.
- [9] Zashita informazii i nadejnost informazionnix sistem "Belorusskiy gosudarstvenniy tekhnologicheskii universitet". Metodicheskiye ukazaniya i kontrolniye zadaniya dlya studentov, Minsk, 2012. https://elib.belstu.by/bitstream/123456789/3372/1/zashhita-informacii-i-nadezhnost_-urbanovich-dlya-z.o.o..pdf.
- [10] K. Ahmad, J. Shekhar, K.P. Yadav. Classification of SQL Injection Attacks. VSRD Technical & Non-Technical Journal Vol. I (4), 2010, pp. 236-242.
- [11] Maraj, G.Jakupi, E.Rogova, Xh.Grajqevci, "Testing of network security systems through DoS attacks", "Mediterranean Conference on Embedded Computing (MECO 2017), IEEE/Scopus conference, Montenegro, June 2017.
- [12] Z. S. Alwan , M. F. Younis. Detection and Prevention of SQL Injection Attack: A Survey. International Journal of Computer Science and Mobile Computing, Vol.6 Issue.8, August- 2017, pp. 5-17.
- [13] J. Clarke. SQL Injection Attacks and Defense. Elsevier, 2012, p.761.
- [14] D.M. Romanenko. Osnovi setevogo administrirovaniya. Minsk: BGTU, 2015. 135 s.