3·4·2021

ЕЛЕКТРОТЕХНИКА И ЕЛЕКТРОНИКА

ELECTROTECHNICA & ELECTRONICA

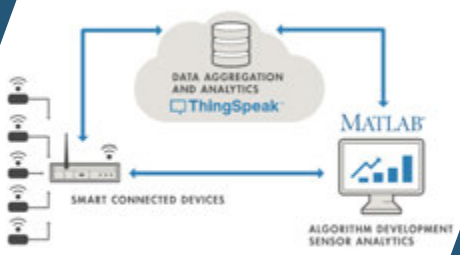
E+E

PUBLISHED BY

BULGARIAN UNION OF ELECTRONICS,

ELECTRICAL ENGINEERING AND

TELECOMMUNICATIONS



CEEC, 108 RAKOVSKI ST, 1000 SOFIA, BULGARIA, CEEC@CEEC.BG, WEB: CEEC.FNTS.BG

ELECTROTECHNICA & ELECTRONICA E+E

Vol. 56. No 3-4/2021

Monthly scientific and technical journal

Published by:

The Union of Electronics, Electrical Engineering and Telecommunications /CEEC/, BULGARIA

Corresponding address:
108 Rakovski St.
Sofia 1000
BULGARIA
Tel. +359 2 987 97 67
e-mail: epluse@ceec.bg
http://epluse.ceec.bg

ISSN 0861-4717

C O N T E N T S

COMPUTER SCIENCE

Krassimira Schwertner
Public and private autonomous data base cloud 29

Ivaylo E. Ivaylov
EEG classification for BCI using genetic algorithm and k-fold cross validation 38

ELECTRICAL ENGINEERING

Mila Ilieva-Obretenova
Impact of an energy conservation measure on reducing CO_2 emissions 46

Public and private autonomous data base cloud

Krassimira Schwertner

The article explains and discusses a new phenomena in organization of the Data Centers. This is known as autonomous database cloud. The private autonomous cloud is the next level in the evolution of the public autonomous cloud services. The properties and functionalities of the private cloud are revealed and explained. The most important features are that the sensitive data is stored on the premises and not in public data centers. This is a big security advantage which is also connected with the control of the places where the standby data is stored. So it is possible to restrict the trans frontier data placement. This often is subject of law restrictions and gives full control over the enterprise data. It is explained why these cloud services should be autonomous, i.e. self-controlled. The price and the limitations of this approach are discussed. No databases that run on-premises or in cloud environments today are 100% autonomous – but that is the goal toward which the industry is headed. The article is based mainly on Oracle Cloud.

Keywords – data center, private autonomous data bases, public autonomous cloud services, oracle 19c

Обществен и частен облак с автономни бази от данни (Красимира Швертнер). *Статията обяснява и обсъжда нови явления в организацията на центровете за данни. Този подход е известен като автономен облак на базата данни. Частният автономен облак е следващото ниво в еволюцията на публичните автономни облачни услуги. Разкриват се и се обясняват свойствата и функционалностите на частния облак. Най-важните характеристики са, че чувствителните данни се съхраняват в предприятията, а не в обществени центрове за данни. Това е голямо предимство за сигурността, което е свързано и с контрола на местата, където се съхраняват данните в режим на готовност. Така че е възможно и да се ограничи трансграничното пренасяне на данни. Всичко това много често е обект на законови ограничения и дава пълен контрол върху данните на предприятието. Обяснено е защо тези облачни услуги трябва да бъдат автономни, т.е.самоуправляеми. Обсъждат се цената и ограниченията на този подход. Нито една база данни, която работи локално или в облачна среда днес, не е 100% автономна - но това е целта, към която се насочва индустрията. Статията се основава главно на разработката на Oracle Cloud.*

1. Introduction

The cloud is a type of computing in which a service provider makes resources available to the public via the internet. Resources vary by provider but may include storage capabilities, applications, Data Bases on virtual machines. Public cloud allows scalability and resource sharing that otherwise would be very expensive or even not possible to afford for a single organization to achieve. Some public cloud providers offer resources for free, while clients pay for other resources by subscription or a pay-per-usage model. Cloud services are available to individual users, as well, and prices scale depending on the user's resource needs. Organizations with huge amounts of data need to develop a cloud migration strategy before choosing a cloud vendor. Private cloud is a computing model that offers a proprietary environment dedicated to a single business entity. As with other types of cloud computing environments, private cloud provides extended, virtualized computing resources via physical components stored on-premises or at a vendor's datacenter.

According to the analysis of Service Express the top challenges for the IT Companies are (Results were gathered between May and October 2020 from over

500 IT professionals in the US) [10]:

- Budget & cost reductions -72%
- New technology and / or adoption – 54%
- Security – 54%
- Talent & workforce – 40%
- Privacy & data compliance – 28%.

Many of these challenges could find resolution in the cloud data center. Obviously 1, 2 and 4 could be achieved using Cloud Services. It is doubtful that 3 and 5 can be guaranteed better in the Cloud. It depends on the quality of the Service.

The analysis of the Top Off-Premises Drivers reveals that the Disaster Recovery/Business Continuity leads with 70%. This is based on the power of the leading Cloud Providers and their capabilities to establish many linked Data Centers with powerful hard- and software. Additionally (it is not mentioned in [10]) the qualitative barrier between big and small companies is broken. Shortly this means that even small companies can use powerful hardware and software business systems on the Cloud which was not affordable for them on the premises because of the immense high costs. Additional Off-Premises Driver are the Maintenance Costs of the on-premises hardware [10]. 56% of the companies complain the High Costs of the Maintenance and also from Administrative Complexity/Delays, inconsistent System Monitoring Quality, and Part Delays. This impose obstacles within every IT budget. The Cloud or the hybrid approach delivers Capital Expenses savings, as well as extends hardware refresh cycles, allowing IT teams to focus on more strategic initiatives.

The article "Comprehensive Analysis of Cloud based Databases" presents various available data storage options, suitability and limitations in context of data science and application of set of data, not relational database [11].

Big Data is a complex multidimensional notion and Big Data analysis is very difficult task. The components of the Big Data are mostly unstructured data. The next problem is the short lifecycle of these data, because data is changed very quickly and replaced by tons of new documents. All these facts leads to the necessity of new technology methods for data processing known as Big Data. Because Big Data comprise an enormous big amount of Data the best solution is to place these Data in Cloud Data Center.

One of the chief advantages of a private cloud deployment is the enhanced degree of control offered to the organization. Because the private cloud is only accessible to a single business, that organization has the ability to configure the environment and manage it in a manner that is uniquely tailored to the specific computing needs of the company.

A private cloud strategy may be comprised of hardware hosted locally at a facility owned by a business, or it may be hosted by a cloud service provider. Virtual private clouds are typically paid for on a rolling basis, but provisioned hardware and storage configurations maintain the benefits of a secure, exclusive network.

2. Public vs. Private Cloud

Common features

There are three general cloud models to choose from, each with a unique set of capabilities and advantages.

A private cloud (also known as an internal cloud or corporate cloud) is the most secure option because the organization has direct control over the infrastructure and only authorized users can access the network. Often the quality of the support is not so good as in the public cloud because of lack of high qualified support personal.

Public cloud services are another popular choice because the enterprise can control costs by reducing on-site hardware investments. With low upfront costs, an organization can deploy an application within the public cloud with ease. Public cloud also allows organizations to fail cheaply if the application does not meet expectations. This can be important for lean businesses that need to reserve capital. The support services in the public cloud are very good because it is possible to hire qualified personal and to develop automatic tools for support.

Hybrid cloud models offer the advantages of public and private clouds by bridging the two models with a layer of proprietary software. Hybrid cloud makes it possible to store vital data in a secure on-site environment while simultaneously leveraging the computing power of the public cloud. Meanwhile, the business only pays for the computing power it uses, allowing for additional cost savings.

In this article, also the private cloud will be discussed. The reason for this is that recently most of the advanced features of the public cloud could be implemented in the private cloud. Up to now this was not possible because the public cloud providers wouldn't like or it was not possible technically to share their tools for automatic maintenance of the hardware and/or software. So the public cloud services were more reliable, the support of the software and the hardware was better and more professional. Another reason that propel the movement of public cloud technologies to private cloud owners was the characteristic of the public cloud provider. There are some cloud services compa-

nies that have closed the product circle - hardware, operating systems, applications and databases. They can deliver the full stack of hardware and software to private cloud customers and so they expect to sell not only cloud solutions but also the needed hardware and software. Even more - the solutions they sell are oriented and can work only on the hardware, operating systems, applications and databases that only they deliver. So the cloud services solution is based on the other products of the company and so increase the value added product of the deal. Other companies provide only cloud services but do not produce and sell many of the other components. They buy them from other suppliers and are dependent from them. So they haven't business reasons to sell cloud technologies to other companies.

One of the main questions standing of front of the private cloud owners is it is worth to purchase cloud technologies based also on the additional need to buy also specific hardware and software. There is no definite answer on this question. Because this business started only in a few month there are not much sources that analyze the pro and contras to this new business product that come to the IT market. The article will try to analyze some of advantages and disadvantages of this approach.

3. Autonomous Cloud

One of the main reason to buy cloud services is that they include autonomous maintenance of the hardware and the software. The autonomy is a very new and important trend in the IT. It is a broad notion, has many aspects and we will concentrate only on the database and application maintenance.

There is a difference between "automatic" and "autonomous" capabilities [4]. A process for database backup, failover or resizing that can be accomplished automatically is still not autonomous if a database administrator has to respond to an alert, make decisions and click a few buttons (or type a few commands) in order to initiate the automated activity. A more dramatic example is when an alert related to a component outage or performance degradation appears automatically on a management console, but doesn't provide sufficient information to diagnose the problem, determine its root cause or offer a definitive recommendation for resolution. The automation literally stops with the alert. What happens next and how long it takes until resolution is unclear.

By contrast, an autonomous database combines the dynamic agility of the cloud with the intelligent responsiveness of applied, adaptive machine learning. The design goal is to minimize or eliminate human labor – and associated human error – and ensure data safety and optimal performance. Businesses will find that autonomous capabilities can further help IT staff improve efficiencies by enabling them to focus on higher value activities in lieu of mundane, time-consuming tasks. This is significant considering that up to 75% of IT budgets are spent on manual database management. An autonomous database can help organizations transform IT operations into a modern cloud model that lowers operating expenses, eliminates costly downtime and ultimately enables them to innovate more while using fewer resources.

The main expectation is that the Oracle Autonomous Database Cloud offers total automation based on machine learning and eliminates human labor, human error, and manual tuning.

The main features are [9]:

- No Human Labor: Database automatically upgrades, patches, and tunes itself while running; automates security updates with no downtime window required.
- No Human Error: SLA guarantees 99.995% reliability and availability, which minimizes costly planned and unplanned downtime to less than 30 minutes a year.
- No Manual Performance Tuning: Database consumes less compute and storage because of machine learning and automatic compression. Combined with lower manual admin costs, Oracle offers even bigger cost savings.

These unprecedented targets at this point of the development of the software sound as pure phantasy for every database administrator (DBA) [8]. This is not possible in the so called on-premises installations, which have the different parameters of the hardware, operating systems and other installation component.

The benefits of the autonomous maintenance are following [3]:

- Automatic Implementation of security features such as detection and protection of privacy-related and other sensitive data, detection of improper or suspicious access, and control of database access in a comprehensive way.
- Automatic Timely patching of database management system (DBMS) software, especially where security patches are involved.
- Protecting Sensitive data and Ensuring Compliance.
- Automatic actions to increase the performance of the software.

3.1 Patching

One of the most time consuming and disruptive for

the applications run is the process of patching of the databases [2].

The main problem that leads to patching is that the computer software is never in a steady state. It requires permanent fixing, development and improvement. Some of this effort has to do with shifting usage models or the correction of previously undetected problems, but a great deal has to do with countering vulnerabilities that may be found and exploited by hackers. This is especially important for databases, where breaches result in significant liability for the enterprise. These updates are applied through patching.

A patch is a piece of code that is inserted into existing software to alter its behavior. It may represent a fix to a known problem, a much-requested enhancement, or the removal of a security vulnerability. Applying a patch to a database server normally requires taking it offline to modify the code and then bringing it up again.

The patching is a big problem. A delay in patching exposes the enterprise to potential hacking. But it is a hard task to do the patch right away. The reason is that patching requires both downtime and staff time, it must be scheduled for off-hours. Even at night, however, taking the database and/or the application offline is going to inconvenience someone, and in the case of a 24 x 7 availability requirement, an offline option is not possible. The alternative is to set up a second database server, load the software, apply the patches, test the patched system, and then swap servers. This approach usually results in brief interruption as one server is quiesced, the last transactions are passed over to the other system, and then the other system comes online. So patching causes extra staff effort, interrupts other work, and disrupts the operations schedule. These activities can represent considerable cost. Most users have many database instances, and the total staff time cost can be calculated at roughly an hour per instance. Even if the patch itself only takes 15 minutes, the process of taking the system offline, applying the patch, verifying the patch, and bringing the system back online can take an hour altogether. We doesn't even take into account the operational disruption as systems administrators and other operations staff must work around this activity. This is why patches are not usually applied as they come in but instead are batched up and applied in bunches at some scheduled time.

The risks of delayed patching are many. When patches are not applied in a timely manner, problems that have been fixed in the current code base are not addressed, improvements are not available and, most importantly, known vulnerabilities are still present, exposing the database to potential hacking. Such delays also create support problems. Very often DBAs get that sinking feeling when they call in with a problem, and the first question the support engineer asks is, "What is your version and patch level?" If it's not current, the advice may be to get up to the current patch level and then call back if the problem persists. This is very bad because it means actually that the enterprise has not support. The security risk is particularly significant. Once patches are issued to address a vulnerability, that vulnerability becomes widely known and hackers are looking for databases that have not yet been patched. So the danger of delayed patching is considerable.

The Cloud Services have faced this problem. That’s why the providers have developed tools to solve it. But this is a complex of measures that include database development over decades and special ad hoc software and can be done only by providers that have “closed the circle”. Moving to the cloud can address the problem of timely patching, yet this is not necessarily a perfect fix. In many cases where a public cloud service is involved, the user may operate under a "shared responsibility model" that places physical responsibility for the system, including its security, in the hands of the managed cloud provider but leaves responsibility for the state of the software, including patching, in the hands of the user. But don’t forget - the patching effort must be done by staff and results in downtime. Even some managed cloud database providers that offer full software support for the database server may need to schedule patching with the user because the patching operation causes an interruption in service. This alone could compel some users to batch up patches rather than have them applied as they come in.

The only way to ensure the timely application of patches is to use a service that offers a nondisruptive patching process. This obviates the need to delay or batch up patches and ensures that the database servers are running with the latest version, including the fixes for all known security vulnerabilities. The critical nature of this capability should be obvious.

The article will discuss the Oracle’s approach to patching. Oracle applies all security patches immediately and on a scheduled basis. Patching is done on a rolling basis, which eliminates downtime, ensuring that the database is continuously available. Oracle can do this because the hardware is configured for nonstop operation, and the software provides a smooth cutover from server to server. Oracle's long history of server clustering for uninterrupted operation — most notably in Real Application Clusters (RAC) introduced in 2001 — has made this possible. The flexibility afforded by the self-tuning functions of the Oracle Autonomous Database software ensures smooth performance throughout. This technology is the result of decades of

research and development at Oracle and first rose to prominence with the self-managing features of Oracle9i. The patching and patch testing processes take place behind the scenes, so the effect is to make the patch process seem virtually undetectable by the user. Contrast this with the manual and error-prone processes of other patching methods. Even more - Oracle can afford upgrading in rolling manner. Upgrade means change the version of the database, e.g. from version 12c to 19c. To achieve this the technologies like Data Guard or Golden Gate can be used.

3.2 Protecting Sensitive data and ensuring Compliance

One part of the sensitive data is relatively easy to identify and protect [2]. These data is formally defined for that purpose under laws, regulations, or contracts and can be easily discovered and defined. This make possible to prevent unauthorized access. Personally Identifiable Information (PII), on the other hand, is harder to detect, especially with the range of emerging privacy regulations, including GDPR, HIPAA, CCPA, and some of other data privacy regulations. In some cases, protecting specific field data is not enough. It is necessary to prevent access to combinations of data that could compromise a person's identity as well. Proper tooling can ensure that those efforts can be done efficiently and consistently.

Sometimes, elements of PII may be embedded in a text field or may be labeled in an obscure or confusing way. Tools that can detect such PII are vital to ensuring that the enterprise remains in compliance, and out of legal trouble.

Obviously, when database applications change, they need to be tested against realistic data — data of the same volume, range of values, and distribution of values. But using real data is out of the question because to do so would enable developers and testers to see it. So a sophisticated data masking approach is called for, one that can produce useful test data that exercises the application just as the real data, but with no elements that could compromise real sensitive data.

Protection from the Cloud Service Vendor is also very important. It is so because any access to sensitive data by an unauthorized person can be considered a violation. Although the cloud service vendor may need to be able to see and manipulate the database operational settings, and even the schema, the vendor should not have access to the data itself.

System-Level Security is the next issue. Managing security at the network, OS, VM, and container level can be tricky at best. Coordinating security definitions across all these levels requires tools that make such settings straightforward to define and modify. While the mechanics of the network, OS, VM, and container-level security are the responsibility of the cloud service provider, how they are set is up to the user. Tools that are inconsistent in operation, or difficult to use, make that job so much harder.

Protection from Malign Access task is important. A constant concern is that of suspicious activity, which may indicate either an external attack or improper access by an internal user. Although there are numerous tools available that can perform database log analysis and identify such activity, it is always up to the data management team to take action, usually well after the fact. What's needed is a database system that can respond immediately when a potential breach is detected while keeping DBAs in the loop. Such detection would include suspicious patterns of access both at the time of access and later (through log analysis, for more sophisticated kinds of breaches).

Oracle Autonomous Database offers five key additional areas of functionality to ensure the security of the database, besides patch management. These are following:

- Encryption. Oracle offers always-on encryption for data at rest and in motion. Data is transferred from storage to processing nodes encrypted. It is even kept encrypted in cache. This function is enabled automatically.
- Separation of the duties of data management and database administration. Oracle is using features such as Database Vault and the Pluggable Database Lockdown profiles to isolate database administration (managed by Oracle) from data administration (managed by the Autonomous Database customer).
- Audit. Data auditing is automatically configured and enabled and in constant operation. The database system records suspicious patterns of access and includes the flexibility to extend analysis of the collected data to other services or even to on-premises security information and event monitoring systems.
- Dedicated infrastructure. Optionally, users may request infrastructure dedicated to the service of their databases only, thereby delivering physical isolation, which is required by some data security regimes.
- Data Safe. Data Safe is a unified database security control center that detects risks introduced by users, data, and configurations through continuous monitoring.

Oracle Data Safe is a comprehensive set of security

features for Oracle Cloud Databases, including Oracle Autonomous Database. It includes the following areas of functionality:

- Security assessment;
- User assessment;
- Activity auditing;
- Data discovery;
- Data masking.

These capabilities are delivered through a single database security control center that allows customers to identify sensitive data and mask it, flags risky users and system configurations, and monitors database activity to quickly discover suspicious attempts to access data. The control center is designed to be easy to use and to save time in performing these critical security tasks. Its design principles include tightly integrated features and reporting, risk dashboards, and extensibility to support new security features. In addition, the control center's security data is physically isolated. Its data masking is designed to provide full protection of PII using over 50 predefined masking formats and a range of mathematically based masking transformations.

3.3 Data Base Machines – maintenance enhancement and performance tuning

A Data Base and Applications Hardware comprise many hardware components. The old school approach is to purchase them from many suppliers and to maintain them from a big staff of technician. This causes many problems - technical, finance, staff and maintenance.

Next issue is that the most important characteristic of a database from user point of view is their performance of the operations and reliability of the data. Performance is a wide topic and field in which many professionals work during his whole career in the IT. It is no surprise that one of the focus of autonomous databases is the performance. It is especially important in the current big databases and data warehouses. To address these needs the industry develops special capsulated database hardware and software complexes known as database machines.

As a sample of Database Machine (DB Machine) it will explain the functionality of Oracle Exadata.

3.3.1 Exadata Database Machine

The Database Machine contains all hardware components needed to run data bases in one cabinet delivered and maintained by one supplier. The list of the components includes Data Base Servers, Storage Servers, and four Switches for internal and external network communication [1]. It is not possible to explain all features of the Exadata, so the article will concentrate on the key features: Data Storage Units - Cells, Smart Flash Cache, Smart scan, Automatic Indexing.

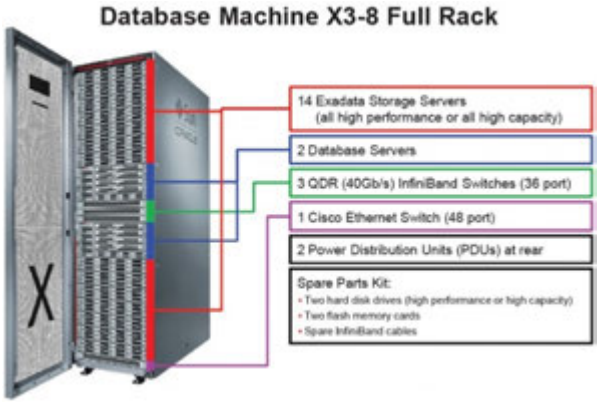


Fig.1. Database Machine X3-8 Full Rack [1].

3.3.2 Data Storage Units - Cells

Data Storage Units - Cells with 12 hard disks every which works as one unit with high degree of intelligence. This intelligence is delivered by 2 Intel Xeon processors included in every cell. The number of cells vary between 3 and 14 in a single appliance [12].

The data storage units are named cells. Each cell comprises 12 identical disks (the type could be different according the needs) and 2 Intel Xeon processors under Linux.

The disk data operations are done on the cells from the software. The SQL statement is sent to the cells using a special protocol iSQL. The cells return only the results of the SQL statement to the DB instance. So the data processing is done on the cell, not on the DB instance minimizing the data transmission. The cells also ensure replication of the data for security reasons (normal redundancy – 2 cells, high redundancy – 3 cells replication). The redundancy helps to keep the instance alive in case of loosed disks in the cells.

The cells implement the Smart Scan feature which does the search of data in the cell and has following advantages:

- Moving SQL processing off the database server frees server CPU cycles and eliminates a massive amount of unproductive I/O transfers. These resources are free to better service other requests. Queries run faster, and more of them can be processed.
- Exadata Storage Server performs join processing for star schemas (between large tables and small lookup tables) in data warehouse systems.

Another useful feature are the Automatic Storage Indexes which are automatically maintained (up to 8 and different over the different parts of the tables) and have following advantages because:

- Business Intelligence (BI) Tools are able dynamically to construct big amount of unpredictable SELECT statements. Practically it is impossible to define in advance the needed indexes to speed up the retrieval.
- Different parts of a same table often is used in different SELECT statements.
- Classic DB indexes use significant disk space.
- By update/insert/delete operation often need for rebalancing of the search tree arises which consumes significant time in the background.

The classic approach is to analyze the application, daily workload and manually to create the needed indexes over the whole table(s). By automated Indexing the DB automatically analyses the workloads and creates the needed indexes. The origins can be found in the old DB machines produced in 2013 that automatically creates up to 8 indexes over parts of the tables for data warehouse processing. If a new index was needed the last recently index was deleted and a new one was made. Nowadays this technology is enhanced significantly and is one of the most attractive feature of Oracle 19c, but firstly - it is paid feature, and secondly - it can be used only on Exadata.

The smart scan and the automatic indexes leverage the performance of the BI tools (analytical processing).

3.3.3 Flash Storage

Exadata Smart Flash Cache (ESFC) focuses on caching frequently accessed data and index blocks, along with performance critical information such as control files and file headers. In addition, DBAs can influence caching priorities using the CELL_FLASH_CACHE storage attribute for specific database objects.

The ESFC is managed by the Exadata Storage Cell Server software **CELLSRV**. In general, when a Database Node requests a block of data from an ASM disk, the CELLSRV software issues asynchronous requests to the ESFC and to the grid disks that underlie the ASM disk group [1]. If the data is in the Flash Cache, this is satisfied from the cache, and if not, from the grid disk. After forwarding the block to the Database Node, CELLSRV then stores any blocks retrieved from the grid disks into the Flash Cache—provided that the blocks are "eligible."

Exadata Smart Flash Cache distinguishes the different types of database I/O [1]:

- Frequently accessed data and index blocks are cached.
- Control file reads and writes are cached.
- File header reads and writes are cached.
- DBA can influence caching priorities.

Smart Flash Cache speeds up the OLTP (On Line Transactional Processing).

All components are controlled and checked by a special monitoring and tracking system. If a malfunction is detected the Oracle support center is informed via automatic generated service request sent to support center by email. This leads to immediate intervention of Oracle technicians which repair or change the defect unit on the fly [5].

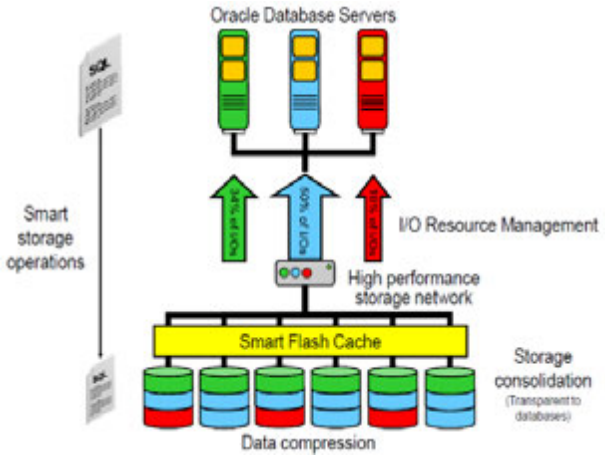


Fig.2. Exadata Cells – Disks, OS, 2 Processors [1].

3.3.4 Advantages of Exadata

The performance of Exadata [6] is very good. This is true for all kinds of data processing paradigms: OLTP, data warehousing, reporting. It is easy to migrate to Exadata via SQL*Loader or Data Pump tools, also to clone the databases.

The support of the hardware and the software is done from only one vendor, so there are no problems with different contracts, guaranties and contacts with different suppliers. The hardware problems are automatically detected and Service Requests to Oracle has been automatically filled.

3.3.5 Disadvantages of Exadata

Exadata has many Linux and Unix operating systems installed in different components and has also running Oracle instances in the Computing Units. Updating and patching these software pieces needs downtime of the systems (4 days in some cases, and rare is done smoothly without collaboration with Oracle Support) [6]. So in order to ensure smoothly running of the applications a prerequisite is to have Data Guard standby instances to that the applications will be switched over. This needs big investment of money to double the appliances. There are also other complaints in regard of the experience of Oracle to support Exadata [6]. Another trouble is that Exadata comes with many licenses on software and this makes it too expensive [7].

4. SWOT Analysis of Private Autonomous Cloud

SWOT is Acronym of Strengths, Weaknesses, Opportunities and Threats. We will concentrate on Oracle proposal for Public and Private Autonomous Cloud because it is a new trend.

4.1 Strengths

The deployment of the Data Bases under the control of experienced professional staff and on highly enhanced hardware is a big advantage. The elastic control over the size of the virtual hardware reduces the loss of investments. The capital investments in Data Center and Hardware are avoided if the company uses a Cloud Data Center. Automatic upgrade without downtime is a big feature for companies that can't tolerate downtimes of its applications. The High Availability based on Data Guard makes the IT processes highly reliable and prevents data loss. The Companies avoid the investments in two data centers located geographically far away of each other to run primary and standby instances according the laws. The usage of Data Base Machines speeds up the Data Processing. High parallelization of disk access speeds up the processing of data. The usage of different kinds of permanent storage of the data makes possible the use Data Lifecycle Management and to keep the different kinds of data on different devices. It is more likely to process the "hot" Data from high speed caches and to keep historical Data on the more slow memory devices. The data base machines make possible to use automatic created partial indexes and dynamically reorganize them. So the old concept of indexing (index all rows in the table, which has many disadvantages and is space consuming) can be replaced by more adaptive indexing methods that are relevant to the contemporary big data volumes. The database machines are equipped with specially designed central processors that are constructed for processing of data and big data like the SIMD (Single Instruction Multiple Data) type of processors.

The private autonomous cloud allows the companies the use of the maintenance software that up to now was used only on the Oracle Cloud.

4.2 Weakness

Every luxury has his price. The advantages listed in S4.1 are the top of the database and application processing but the costs of their usage is enormous. The hidden trap are the database machines. Many of the features explained above can be run only on the special designed database machines which are expensive. The null downtime update and upgrade of the software is only possible if there is a second standby database machine which is running the processing while the upgrades/updates run on the first machine. So practically the hardware is doubled. The licenses fees are also doubled. The database machines are very intelligent, but this intelligence is based on special software installed on special processors inside the machines. This software needs to be regularly upgraded/updated which not an easy task is. Of course the same activities should be planned in the normal data centers, but the complexity of the tasks there is not so high. Additionally the transmission the data to the data centers and back to the customers is also paid on volumes basis. There are some problems with the scaling the database machines. In some cases it is not possible to enlarge physically the machine because all slots are occupied. Many of the functionalities (like automated indexing in Oracle 19c) are additionally paid features and are available only on the data base machines. The next danger is the theft of the corporate data. The data in motion (transmitted via Internet) can be captured by hackers, the data in rest (stored on disk devices in the data centers) can be captured by internal personal. The next legal problem is that by crashes of any kind standby data centers are using to continue the processing. Often these data centers are placed in different countries. This means that copies of the corporate data is placed in different countries and in some special cases this is prohibited by the local low. This is true for the public cloud. Partially the two last problems can mitigated by using private autonomous cloud. Unexpectedly the internet connectivity and throughout grow in importance and play vital role to insure the operating capability of the industry and the business.

4.3 Opportunities

The implementation of Applications and Corporate Databases will be possible not only for elite Enterprises. Middle and small Companies will be able to use the Cloud Services to base their businesses on the best applications and databases and so will enhance significantly the quality of the IT services in the companies. The big and the middle enterprises can establish their own Data centers which will be able to use autonomy of application and database maintenance and will make usage of the enhancement of the data processing delivered from the database machines.

4.4 Threats

The concentration of Data Processing Services in Cloud Data Centers makes the functionality of the enterprises independent of the connection lines (Internet), the possible disasters in the data centers (electricity, connections and network, fire, flood, physical attacks, hacker attacks with massive global effect).

Conclusion

The concentration of massive hardware and software resources in centralized Data Centers allows agile end highly elastic usage of the available resources, savings of capital investments in buildings and hardware. It makes possible for small and medium businesses to use expensive software applications and databases on affordable cost what was impossible in the past. The usage of specialized database machines increases the databases performance and enhance the analytical processing in data warehouses. The new trend is autonomous maintenance of the databases. The next steps is to allow the enterprises to establish their private autonomous cloud centers which is additional business niche for the big hardware and software producers.

Based on our long experience as DBA we analyze as deep as possible the new trends and their components. The challenge was to distinguish the marketing claims from the real price of the technologies and to reveal hidden costs and efforts behind the new technologies.

REFERENCES

[1] Oracle Corporation, A Technical Overview of the Oracle Exadata Database Machine and Exadata Storage Server https://www.oracle.com/technetwork/database/exadata/exadata-dbmachine-x4-twp-2076451.pdf

[2] Olofson C. The Security Benefits of a Fully Managed Database Service: Oracle Autonomous Database, IDC.

[3] Oracle Corporation, Oracle Autonomous Database Strategy, July, 2020.

[4] Oracle Autonomous Database, White Paper Series: The Industry's First Self-Driving Database, September 2018.

[5] Hall T., Oracle Database 18c Released: How to Get Started, https://oracle-base.com/blog/2018/02/17/oracle-database-18c-released-how-to-get-started/.

[6] Mason J. Oracle Database Appliance: The Good, The Bad and the Ugly, http://houseofbrick.com/oracle-database-appliance-the-good-the-bad-and-the-ugly/, January 20, 2018

[7] Kanaracus C. Experts: Avoid Big Mistakes with Oracle's Exadata, https://www.cio.com/article/2376980/enterprise-software/experts--avoid-big-mistakes-with-oracle-s-exadata.html, Apr 18, 2014.

[8] Hall T. Oracle Database 18c is NOT an Autonomous Database! https://oracle-base.com/blog/2018/01/03/oracle-database-18c-is-not-an-autonomous-database/, January 3, 2018.

[9] Oracle Corporation, The World's #1 Database Is Now the World's First Self-Driving Database, https://www.oracle.com/database/autonomous-data-base/feature.html.

[10] Data center & Infrastructure Report, Service Express, https://serviceexpress.com/resources/articles-and-insights/it-data-center-infrastructure-report-2021/, 2021.

[11] Kumar, E. Saravana, Selvaraj Kesavan, R. Ch. A Naidu, Senthil Kumar R, Latha. Comprehensive Analysis of Cloud based Databases, ICETCE, 2021, https://iopscience.iop.org/article/10.1088/1757-899X/1131/1/012021/pdf.

[12] Schwertner K. Towards autonomous hardware and software in data base management systems, Journal Electrotechnica & Electronica (E+E), vol.52, No 11-12, pp.20-26, 2017, ISSN 0861-4717.

Prof. Krassimira Schwertner (PhD) *works in Sofia University. Her research is focused on databases, database-as-a-service, database security, database in-memory, and engineered database appliances. She has published more than 110 papers.*

e-mail: shvertner@feb.uni-sofia.bg

Received on: 26.04.2021

EEG classification for BCI using genetic algorithm and k-fold cross validation

Ivaylo E. Ivaylov

Brain-Computer Interfaces (BCIs) have become popular in the last years and are used for several applications in human-computer interfaces including attention and workload measures as well as for direct control of objects. In this paper an electroencephalography (EEG) signal classification approach is applied for BCI. In this paper an approach based on genetic algorithm and k-fold cross validation is suggested for the EEG signal de-noising problem in motor imagery BCI. In order to evaluate the influence of the proposed feature extraction and feature selection approach on the accuracy of EEG signal classification five classification algorithms are adopted that are selected based on their performance in solving signal processing problems: Logistic Regression, K-neighbors, Support Vector Classifier (SVC) with Linear Regression, SVC with Radial Basis Function (RBF) Regression and Gaussian Training Classifier. The experimental results on Electrocorticography (ECoG) mental dataset for motor imagery BCI show that the classification accuracy can be improved by use of the suggested utilization of and k-fold cross validation at the feature extraction stage and genetic algorithm at the feature selection stage and the additional computational overhead is not significant.

Keywords – brain computer interface, electro-encephalography, feature extraction, feature selection, genetic algorithm, k-fold cross validation

BCI класификации, използващи генетичен алгоритъм и k-fold кръстосана валидация (Ивайло Е. Ивайлов). *Мозъчно-компютърните интерфейси стават все по-популярни през последните години. Изследователите използват тази технология за няколко типа приложения, включително приложения за привличане на внимание и натоварване, но също така и за директен контрол на обекти с помощта на BCI. В статията е представен подход за класификация на електроенцефалографски сигнали, базиран на използване на генетични алгоритми и k-кратна валидация за извличане и избор на признаци с цел подобряване на точността на класификацията. За да се оцени влиянието на предложеният подход върху точността на класификацията на ЕЕГ сигнали са използвани пет класификационни алгоритма, избрани поради ефективност им при обработка на сигнали: когистична регресия, класификация с K-съседи, класификация с опорни вектори и линейна регресия, класификация с опорни вектори и регресия с радиална базисна функция и класификатор на Гаусово обучение. Експерименталните резултати, получени за обучаващи електрокортикографски (ECoG) ментални данни показват, че точността на класификацията може да се подобри чрез използване на предложения подход с използване и k-кратно кръстосано валидиране на етапа на извличане на характеристика и генетичен алгоритъм на етапа на избор на характеристика, при това допълнителните изчислителни ресурси са несъществени спрямо подорението на точнсотта на класификацията.*

Lis of used abbreviations

ABC	Artificial Bee Colony
ACO	Ant Colony Optimization
BCI	Brain-Computer Interface
BPNN	Back-Propagation Neural Networks
BPSO	Binary Particle Swarm Optimization
BPSO-CSP	Binary Particle Swarm Optimization-Common Spatial Pattern
COP	Constraint Optimization Problems
CSP	Common Spatial Pattern
ECoG	Electrocorticography
EEG	Electroencephalography
FBCSP	Filter Bank Common Spatial Pattern
GA	Genetic Algorithm
GNMM	Genetic Neural Mathematic Method

ISAGA	Improved Simulated Annealing Genetic Algorithm
LDA	Latent Dirichlet Allocation
MCSP	Multi-channel Common Spatial Pattern
MLDW-PSO	Multi-stage Linearly-Decreasing Inertia Weight-Particle Swarm Optimization
MLP	Multilayer Perceptron
MLP-ABC	Multilayer Perceptron-Artificial Bee Colony
MOPSO	Multi-Objective Particle Swarm Optimization
PSD	Power Spectrum Density
PSO	Particle Swarm Optimization
RBF	Radial Basis Function
SVC	Support Vector Classifier
SVM	Support Vector Machine
WT	Wavelet Transform

Introduction

A brain-computer interface (BCI) can be characterized as a framework that interprets mind action examples of a human and translates them into messages or orders for an intuitive application in human-computer interfaces [1]. A human cerebrum action is estimated utilizing electro-encephalography (EEG) signals and motor imagery actions are used for different machine control operations. For example, an EEG based BCI can empower a human to move a cursor on a monitor by envisioning left or right hand movements. As they make computer control conceivable with no actual activity, EEG-based BCIs guarantee to alter numerous applications' zones and thus remarkably empower seriously motor-impaired humans and allow them to control assistive innovations, for example text input frameworks or wheelchairs [1]. They can be also utilized as recovery gadgets for stroke patients as new gaming input gadgets or used to plan versatile human-computer interfaces that can respond to the human psychological states.

To utilize a BCI, two stages are usually required: (1) an offline preparing stage during which the framework is aligned and (2) an operational online stage wherein the framework can perceive mind action designs and make an interpretation of them into machine commands. An online BCI framework involves a closed loop that start with the human providing a particular EEG signal (for example utilizing engine symbolism) and these EEG signals being estimated. At that point, EEG signals are normally preprocessed utilizing different spatial channels, outliers are removed from the signal and features are extracted. The extracted EEG features are then grouped prior to being converted into an order for a BCI application and before input is given to the humans to educate them if a particular mental order is perceived [1].

A metaheuristic frameworks are widely used approach for mathematical optimization that can be applied to discover, create, select and apply heuristics that may give an adequately decent answer for an optimization problem, particularly in cases of fragmented or uncertain data are available or when a calculation limit should be applied. Metaheuristics are successfully applied to hard optimization problems in cases the search space of the possible solutions doesn't allow exhaustive search to be applied. By utilizing guided heuristic search of the solution space the metaheuristics algorithms are very useful for combinatorial optimization problems [1].

Compared with other optimization techniques and approaches, metaheuristics do not ensure that a global optimal solution will be found but rather provide near-optimal solution for a limited computational time. Various metaheuristic algorithms employ different stochastic improvements and thus the solution provided is subject to randomness applied in the optimization procedure [1]. In combinatorial optimization problems through exploration of huge number of possible solutions metaheuristics can regularly discover good solution of the problem in hand with less computational efforts than improvement based approaches, iterative techniques, or basic heuristics. Metaheuristic-based approaches are usually classified into two main categories based on the number of solutions that are improved during the heuristic search: population based that utilize swarm intelligence approaches for heuristic improvement of many solutions in search of the optimal one, and trajectory-based algorithms that focus on improvement of a single solution by adopting heuristic based local search techniques. Genetic algorithms (GA) are one of the widely used metaheuristics for solving optimization problems that are based on the natural evolution process and exploit "survival of the fittest" principle.

Both feature selection and classification stages are very important and influence the results of the successful application of EEG signals in BCI [1]. Most brain-computer interfaces (BCIs) use the electroencephalogram (EEG) to measure brain activity. Alternatively, the electrocorticogram (ECoG) can be used, which provides better signal quality, but requires the implantation of subdural electrodes. Considering recent advances in signal processing, one might argue that employing modern spatial filters that can

considerably improve signal quality and therefore the utility of EEG (as compared to ECoG), renders the ECoG unnecessary for BCIs [2]. In this paper an approach based on k-fold cross validation and genetic algorithm is suggested for the EEG signal de-noising problem in motor imagery BCI. In order to evaluate the influence of the proposed feature extraction and feature selection approach on the accuracy of EEG signal classification five classification algorithms are adopted that are selected based on their performance in solving signal and image processing problems.

Related work

Various researchers have used metaheuristics in order to improve the results of BCI applications. In [5] genetic algorithm is applied to find the optimum combination of feature extraction methods and classifiers in BCI applications. GA is used in [6] for automatic feature extraction in P300 signal detection problem applied for BCI. Integration of fuzzy classification with particle swarm optimization (PSO) is utilized in [7]. In [8] a novel technique for feature selection in BCI based on PSO and hybrid PSO is suggested. Heuristic algorithms are used in [9] to find a reasonable combination of weights and biases as well as of an appropriate architecture of multi-layered perceptron (MLP) neural networks utilized for BCI classification. A hybrid PSO-GA approach is used for feature selection in [10]. In an improved simulated annealing genetic algorithm (ISAGA) approach is used for feature selection of EEG signal. Incremental quantum PSO optimization algorithm is suggested in [12] for incremental classification of EEG data stream. GA is applied in [13] for feature reduction of EEG signals in mental task classification. Artificial Bee Colony (ABC) metaheuristics is used to select features from the BCI feature combination in [14] and the selected sub-features are classified by Support Vector Machine (SVM). In [15] binary PSO (BPSO) based channel selection of EEG signals is applied and its application to speller systems is presented. Common Spatial Pattern (CSP) and BPSO are used in combination to select the best frequency band because of their high performances in feature extraction of EEG signals and evolutional search respectively. In the BPSO-CSP algorithm each particle consists of 10 components, corresponding to one sub-band of the broad frequency band. Thereby each particle represents a combination of selected sub-bands and thus is a potential solution to the frequency band selection. Depending on the number solution representation one or several sub-bands may be selected by each particle. The covariance matrices of filtered EEG data are first calculated in each chosen sub-band and then aggregated. The CSP algorithm is then applied to the aggregated covariance matrix in order to extract spatial features. Finally, LDA classifier is used for the extracted features. In [17] genetic neural mathematic method (GNMM) method is applied to the EEG channel selection and classification problems. In [18] P300 signals are detected by employing Power Spectrum Density (PSD) for feature extraction and MLP-ABC scheme as a classifier. Artificial neural networks and PSO designed in a hybrid structure are used in [19] for diagnosis of epilepsy patients via EEG signals. Five metaheuristic algorithms are adopted to find the optimal wavelet transform (WT) parameters for the EEG signal denoising problem in [3]. In [20] a modified PSO with multi-stage linearly-decreasing inertia weight strategy (MLDW-PSO) based feature selection is proposed to select the most informative weighted features in an effective way. Swarm intelligence algorithms are suggested in [21] as a reliable method for the optimization of EEG signals for the improvement of the performance of the brain interfaces based on stable states visual events. A novel signal processing stage comprised of Filter Bank Common Spatial Pattern (FBCSP) for feature extraction, PSO for feature selection and LDA for classification is implemented as part of a BCI system in [22]. Walsh-Hadamard transform is applied for feature extraction and feature selection is based on BPSO in [23] and the feature classification uses MLP with back propagation training algorithm and Levenberg-Marquardt training algorithm. In [24] a novel application of a multi-objective particle swarm optimization (MOPSO) is suggested to solve the problem of effective channel selection for BCI systems. Multi-channel common spatial pattern (MCSP) is proposed to extract the features in [25] and two novel channel selection approaches are used to compromise between the optimal channels number and classification accuracy. A methodology for the automatic detection of normal, epilepsy and brain death from recorded EEG signals collected from clinic is proposed in [26]. Discrete WT is applied for feature extraction and error back propagation NN optimized by PSO is used for classification of neurological disorders. Simple BPNN has several drawbacks which mainly include large time duration during EEG signal classification that is removed by the utilization of PSO. A mental task classification algorithm using hybrid approach with PSO and recurrent neural network is suggested in [27]. Features are extracted from EEG signals that are recorded during five mental tasks, namely baseline–resting, mathematical multiplication, geometric figure rotation, letter composing and visual counting. The features are used by the neural network to classify different combinations of two mental tasks. The output of the BCI could be used with some translation schemes like Morse code or as two way movement control of a device as well as serve as

communication or control channel for paralyzed patients with motor impairments. In [28] an optimization based on PSO, Ant Colony Optimization (ACO), GA and differential evolution algorithms is used in order to generate an optimum subset of features that improves the identification of features of EEG signals. Spectral Density of Power, Spectral Coherence methods and the computational cost between these algorithms are presented as measure of comparison.

EEG classification using genetic algorithm and k-fold cross-validation

In this paper an approach for BCI classification of EEG data using genetic algorithm and k-fold cross-validation is suggested. The general EEG data processing pipeline is shown on fig. 1. At the first stage feature extraction of the raw EEG data is based on k-fold cross validation. The second stage uses GA for feature selection and finally at the third stage five classification algorithms are used and experimentally evaluated: Logistic Regression, K-neighbors, Support Vector Classifier (SVC) with Linear Regression, SVC with Radial Basis Function (RBF) Regression and Gaussian Training Classifier. The classification algorithms are carefully selected based on their good performance in solving various signal and image processing problems [3].

K-fold cross validation for feature extraction

Cross-validation is a resampling method used to assess models on a restricted information test. The approach uses single parameter k referring to the number of groups that a given data sample is to be split into.

Cross-validation is primarily used in applied machine learning to estimate the accuracy of a machine learning model on unseen data. That is to use a limited sample in order to estimate how the model is expected to perform in general when used to make predictions on data not used during the training of the model. The popularity of k-fold cross validation is due to it is simple and generally results are less biased or less optimistic estimate of the model is given than other approaches as simple train/test split.

The general k-fold cross validation procedure is shown in Algorithm 1.

At the first step each observation in the data sample is assigned to an individual group and stays in that group for the duration of the procedure thus each sample is used in the hold out set 1 time and used to train the model k–1 times.

Algorithm 1 K-fold cross validation pseudo-code

1: Shuffle the dataset randomly
2: Split the dataset into k groups
3: **for** each unique group:
4: Take the group as a hold out or test data set
5: Take the remaining groups as a training data set
6: Fit a model on the training set and evaluate it on the test set
7: Retain the evaluation score and discard the model
8: Summarize the skill of the model using the sample of model evaluation scores

In order to apply k-fold cross validation at the feature extraction stage of EEG classification the input EEG data signal is segmented into small overlapping regions and then randomly split into training and validation sets.

Genetic algorithm for feature selection

GA is a population based metaheuristics based on the metaphor of the natural phenomenon of Darwin evolution theory.

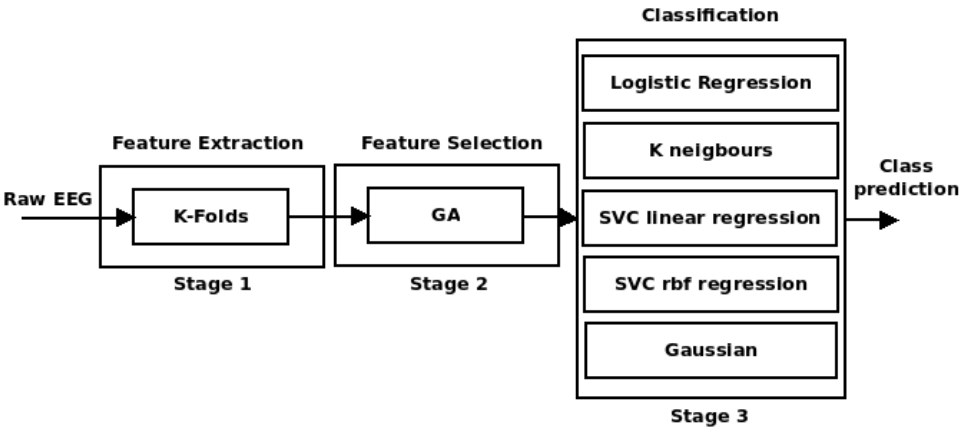


Fig. 1. Data processing pipeline.

GA procedure starts with a set of many solutions of the problem in hand with each solution being a vector of decision variables and each decision variable having a specific range of values. In the evolution context the set of solutions is equivalent to a population, each solution is analogous to a chromosome, each decision variable is analogous to a gene and each value of the decision variables is analogous to an allele [3].

In order to apply GA to constraint optimization problems (COP) both the objective function and the problem representation must be properly adjusted together with parameter tuning. GA typically has a set of parameters, including the size of the population P_{size}, the number of generations P_{no}, the crossover rate $P_{crossover}$ and the mutation rate $P_{mutation}$. In order to build an efficient and robust GA the parameter settings of each COP must be closely examined.

Algorithm 2 shows the high-level schematic pseudo-code of GA that starts with a population of candidate solutions X_{chrom}, where X_{chrom} is an augmented matrix of size $P_{size} \times N$ and N is the number of decision variables in each solution.

Algorithm 2 Genetic Algorithm pseudo-code

1: $X_{chrom} \leftarrow$ Generate Initial Population
2: Evaluate(X_{chrom})
3: **while** (Stopping criterion is not met) **do**
4: $X_{chrom'} \leftarrow$ Selection(X_{chrom})
5: $X_{chrom''} \leftarrow$ Crossover($X_{chrom'}$)
6: $X_{chrom'''} \leftarrow$ Mutation($X_{chrom''}$)
7: Evaluate($X_{chrom'''}$)
8: $X_{chrom} \leftarrow$ Replacement($X_{chrom'''} \cup X_{chrom}$)
9: **end while**

Initially, the population X_{chrom} is filled with random candidate solutions across the problem search space, that is:

(1) $$X_{chrom} = X_{chrom^1}, X_{chrom^2}, \dots, X_{chrom^{Psize}}$$

Each candidate solution $X_{chrom'}$ is evaluated based on an objective function. The improvement loop in GA (Algorithm 1, line 3 to 9) repeats the following steps until a termination criterion is met: select the parents (new population $X_{chrom'}$) that will be used to generate the next population which will pairwise crossover with a probability of $P_{crossover}$ to come up with a new population $X_{chrom''}$. Afterward, each pairwise solution will be checked if it must be mutated with probability $P_{mutation}$ to come up with $X_{chrom'''}$. The new population will be reevaluated and the $X_{chrom'''}$ will be substituted with the population X_{chrom} based on a selection method. This procedure is followed to determine whether the offspring are fit or not. The GA procedure is repeated several times until an optimal solution is reached [3].

In order to apply GA at the feature selection stage of the EEG data classification, the data obtained after the feature extraction stage are represented as an individual with chromosomes that corresponds to the EEG data.

Experimental results

In order to experimentally evaluate the performance of the suggested approach for EEG data classification using 2-fold cross validation at the feature selection stage and genetic algorithm at the feature selection stage with five different classification algorithms EEG data for 2-class motor imagery classification is based on BCI competition III dataset I [29]. The dataset is provided by University of Tübingen, Germany, Institute of Medical Psychology and Behavioral Neurobiology, Max-Planck-Institute for Biological Cybernetics, Tübingen, Germany and Universität Bonn, Germany and comprises ECoG recordings from the same subject and with the same task but as two different sessions on two different days with about one week in between [30]. During the BCI experiment a subject had to perform imagined movements of either the left small finger or the tongue. The time series of the electrical brain activity is picked up during these trials using a 8x8 ECoG platinum electrode grid which is placed on the contralateral (right) motor cortex. All recordings are performed with a sampling rate of 1000 Hz and after amplification the recorded potentials are stored as microvolt values. Every trial consists of either an imagined tongue or an imagined finger movement and is recorded for 3 seconds duration. To avoid visually evoked potentials being reflected by the data the recording intervals start at 0.5 seconds after the visual cue has ended. A labeled training data set of ECoG recordings from the first session consists of two parts: (i) the brain activity during 278 trials stored in a 3D matrix named using the following format: [trials x electrode channels x samples of time series], and (ii) the labels of the 278 trials stored as a vector of –1/1 values.

For the experimental evaluation of the EEG mental data classification, the implementation of the data processing pipeline is based on Python 3.8.3 and the experiments are carried out using Jupyter notebook under Linux Fedora 33, 64 bit, kernel 5.9 on a machine equipped with Intel 12X Core I7-8750H, 2.20 GHZ, 16 GB RAM. Due to large size of the dataset the experimental evaluation is based on 25% of the data.

The accuracy of the classification is measured using several metrics: precision, recall, f1-score and support:

- Precision is calculated as ratio of system generated results that correctly predicted positive observations (True Positives) to the system's total predicted positive observations, both correct (True Positives) and incorrect (False Positives):

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

- Recall is calculated as ratio of system generated results that correctly predicted positive observations (True Positives) to all observations in the actual malignant class (Actual Positives):

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

- F1-score is the weighted average of the precision and the recall and therefore this score takes both False Positives and False Negatives into account to strike a balance between precision and recall:

(4) $$F1 - score = \frac{2*(Recall*Precision)}{Recall+Precision}$$

- Support is the number of samples of the true response that lie in the relevant class.
- Accuracy is the most intuitive performance measure calculated as ratio of the true classified observations over the total number of observations:

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

- Macro average metrics is calculated for each label as unweighted mean.
- Weighted average is the average weighted by support of each label, i.e. the number of true instances for each label.

In order to apply GA at the feature selection stage of the EEG data classification the following parameters are used: population size 300, number of generations 1000, mutation rate 0.5.

The results of the classification accuracy for the EEG data classification with or without utilization of GA and 2-fold cross validation as feature detector and feature selector are given in Table 1 and on Fig. 3.

The execution time of the EEG classification using the five selected classifiers with and without utilization of GA and 2-fold cross validation is shown in Table 2.

The results from the experimental evaluation on the test set for the K-neighbors classifier are given in Table 3 while the results of the SVC RBF Regression classifier on the test dataset using GA and 2-fold cross validation give precision 0.72 and 0.60 for the two labels.

The comparison of the results show that the use of GA for feature selection improves the classification accuracy and the improvement is between 2 to 9 % for each of classification algorithms. The most significant influence of GA based feature selection is observed for the Logistic Regression, SVC Linear Regression and Gaussian Classification, but the overall classification accuracy of Gaussian Classification is quite low compared to the other classification algorithms. The best accuracy is achieved using SVC Linear Regression and Logistic Regression classifiers with GA based feature selection, 0.99 and 0.94 respectively.

On the other side, a comparison of the execution time of each of the classifiers is given in Table. 2. The results show that for all of the five classification techniques the use of GA and 2-fold cross validation requires more time, as can be expected, and the slowdown of the processing is between 1 and 22%, the smallest for K-neighbors classifier and the biggest for Logistic Regression. The fastest processing time is measured using SVC RBF Regression classifier and K-neighbors is the slowest classification technique used for EEG data classification. The best overall performance in terms of both accuracy and execution time is observed for SVC Linear Regression classifier.

Fig. 3. Classification accuracy with and without feature selection using Genetic Algorithm.

Table 1

Accuracy of the classification with and without feature selection using Genetic Algorithm

Classifier	Training Accuracy	
	without GA	with GA
Logistic Regression	0.86	0.94
K-neighbours	0.76	0.78
SVC Linear Regression	0.91	0.99
SVC RBF Regression	0.78	0.82
Gaussian Classification	0.50	0.59

Table 2

Execution time of the EEG classification

Classifier	Execution time [ms]	
	without GA	with GA
Logistic Regression	51.4	66.1
K-neighbours	771	779
SVC Linear Regression	2.48	2.65
SVC RBF Regression	1.06	1.09
Gaussian Classification	14.1	16.4

Table 3

Accuracy metrics

	Precision	Recall	F1-score	Support
Label (–1)	0.69	0.61	0.65	134
Label (+1)	0.49	0.57	0.53	87
Accuracy			0.60	221
Macro avg	0.59	0.59	0.59	221
Weighted avg	0.61	0.60	0.60	221

Conclusion

In this paper an approach based on genetic algorithm and 2-fold cross validation is suggested for the EEG signal de-noising problem in motor imagery BCI. In order to evaluate the influence of the proposed feature extraction and feature selection approach on the accuracy of EEG signal classification five classification algorithms are experimentally evaluated: Logistic Regression, K-neighbors, Support Vector Classifier (SVC) with Linear Regression, SVC with Radial Basis Function (RBF) Regression and Gaussian Training Classifier. The experimental results show that the classification accuracy can be improved by use of the suggested utilization of 2-fold cross validation at the feature extraction stage and genetic algorithm at the feature selection stage and the additional computational overhead is not significant. EEG data processing based on 2-fold cross validation, genetic algorithm and SVC Linear Regression classifier provides best training and test accuracy improving the accuracy by 8% with only 6% overhead of the execution time.

As a future work more robust feature extraction approach can be also adopted to handle data de-noising in the EEG data processing pipeline for motor imagery BCI. Additionally, the suggested processing model can be also applied on different data sessions to assess its ability to generalize to other data.

Acknowledgements

The work presented in the paper is supported by research grant 202ПД0001-09 financed by the Research and Development Sector of TU-Sofia.

REFERENCES

[1] Lotte, F., L. Bougrain, A. Cichocki, M. Clerc, MCongedo, A. Rakotomamonjy, F. Yger. A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. Journal of Neural Engineering, Volume 15, Number 3, 2018.

[2] Graimann B., G. Townsend, J.E. Huggins, S.P. Levine, G. Pfurtscheller. A Comparison between using ECoG and EEG for Direct Brain Communication, July 2005.

[3] Alyasseri, Z., A. Khader, M. Al-Betar, S. Makhadmeh. A. Abasi. EEG signals denoising using optimal wavelet transform hybridized with efficient metaheuristic methods. IEEE Access, Vol. 8, 2020, pp. 10584-10605.

[4] Baig, M., N. Aslam, H. Shum, L. Zhang. Differential evolution algorithm as a tool for optimal feature subset selection in motor imagery EEG. Expert Systems with Applications, Vol. 90, 2017, pp. 184-195.

[5] Eslahi, S., N. Dabanloo, K. Maghooli. A GA-based feature selection of the EEG signals by classification evaluation: Application in BCI systems. ArXiv, abs/1903.02081, 2019.

[6] Dal Seno, B., M. Matteucci, L. Mainardi. A genetic algorithm for automatic feature extraction in P300 detection. Proc. of IEEE International Joint Conference on Neural Networks, Hong Kong, China, 2008, pp. 3145-3152.

[7] Wu. S.-L., et. al. A motor imagery based brain-computer interface system via swarm-optimized fuzzy integral and its application. Proc. of IEEE International Conference on Fuzzy Systems, Vancouver, BC, Canada, 2016, pp. 2495-2500.

[8] Akilandeswari, K., G. Nasira. A hybrid meta heuristic feature selection algorithm for brain computer interface. International Journal of Scientific & Engineering Research, Volume 6, No. 5, 2015.

[9] Afrakhteh, S., M.-R. Mosavi, M. Khishe, A. Ayatollahi. Accurate classification of EEG signals using

neural networks trained by hybrid population-physic-based algorithm. International Journal of Automation and Computing, Vol. 17, No. 1, 2020, pp. 108-122.

[10] Sridhar, G., M. Murali, A. Vamsidhar. An hybrid particle swarm optimization with multilayer perceptron for EEG classification. International Journal of Management, Technology and Engineering, Vol. 8, No. 12, 2018, pp. 5797-5805.

[11] Ji, Y., X. Bu, J. Sun, Z. Liu. An improved simulated annealing genetic algorithm of EEG feature selection in sleep stage. Proc. of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, Jeju, South Korea, 2016, pp. 1-4.

[12] Hassani, K., W. Lee. An incremental framework for classification of EEG signals using quantum particle swarm optimization. Proc. of IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications, Ottawa, ON, Canada, 2014, pp. 40-45.

[13] Rezaee, A. Applying genetic algorithm to EEG signals for feature reduction in mental task classification. International Journal of Smart Electrical Engineering, Vol. 5, No. 1, 2016, pp. 1-4.

[14] Hsu, W.-Y., Y.-P. Hu. Artificial bee colony algorithm for single-trial electroencephalogram analysis. Clinical EEG and Neuroscience, Vol. 46, No. 2, 2015, pp. 119-125.

[15] Arican, M., K. Polat. Binary particle swarm optimization (BPSO) based channel selection in the EEG signals and its application to speller systems. Journal of Artificial Intelligence and Systems, Vol. 2, 2020, pp. 27-37.

[16] Wei, Q., Z. Weid. Binary particle swarm optimization for frequency band selection in motor imagery based brain-computer interfaces. Journal Bio-Medical Materials and Engineering, Vol. 26, No. S1, 2015, pp. S1523-S1532.

[17] Yang, J., H. Singh, E. Hines, F. Schlaghecken, D. Iliescu, M. Leeson, N. Leeson. Channel selection and classification of electroencephalogram signals: an artificial neural network and genetic algorithm-based approach. Artificial Intelligence in Medicine, Vol. 55, No. 2, 2012, pp. 117-126.

[18] Aytekin, S., T. Kıyan. Detection of P300 based on artificial bee colony, Proc. of the International Joint Conference on Biomedical Engineering Systems and Technologies, 2016, pp. 183–189.

[19] Arslan, S., G. Tezel, H. Islk. EEG signals classification using a hybrid structure of ANN and PSO. International Journal of Future Computer and Communication, Vol.1, No. 2, 2012, pp. 170-172.

[20] Li. Z, L. Qiu, R. Li, Z. He, J. Xiao, Y. Liang, F. Wang, J. Pan. Enhancing BCI-based emotion recognition using an improved particle swarm optimization for feature selection, Sensors, Vol. 20, No. 11, 2020.

[21] Fernandez-Fraga, S., M. Aceves-Fernandez, J. Pedraza-Ortega, S. Tovar-Arriaga. Feature extraction of EEG signal upon BCI systems based on Steady-state visual evoked potentials using the ant colony optimization algorithm. Discrete Dynamics in Nature and Society, Vol. 2018.

[22] Cantillo-Negrete, J., R. Carino-Escobar, P. Carrillo-Mora, J. Gutierrez-Martinez. Increasing stroke patients motor imagery classification by selecting features with particle swarm optimization. Proc. of 7th Graz Brain-Computer Interface Conference, 2017, pp. 52-57.

[23] Akilandeswari, K., G. Nasira. Multi-layer perceptron neural network classifier with binary particle swarm optimization based feature selection for brain-computer interfaces. International Journal of Computer and Information Engineering, Vol. 9, No. 6, 2015.

[24] Hasan, B., J. Gan. Multi-objective particle swarm optimization for channel selection in brain-computer interfaces. School of Computer Science and Electronic Engineering, University of Essex, 2017.

[25] Sun, H., Y. Zhang, B. Gluckman, X. Zhong, X. Zhang. Optimal-channel selection algorithms in mental tasks based brain-computer interface. Proc. of 8th International Conference on Bioscience, Biochemistry and Bioinformatics, ACM, New York, NY, USA, pp. 118–123.

[26] Kshirsagar, P., S. Akojwar. Optimization of BPNN parameters using PSO for EEG signals, Advances in Intelligent Systems Research, Vol. 137, 2016, pp. 385-394.

[27] Hema, C., M. Paulraj, S. Yaacob, A. Adom, R. Nagarajan. Particle swarm optimization neural network based classification of mental tasks. Proc. of 4th Kuala Lumpur International Conference on Biomedical Engineering, Vol 21, Springer, Berlin, Heidelberg, 2008, pp. 883-888.

[28] Fernandez-Fraga S., M. Aceves-Fernandez, J. Soto, J. Arreguín. Population-based algorithms applied to brain-computer interfaces upon steady-state visual evoked potentials, Transactions on Machine Learning and Artificial Intelligence, Vol. 7, No. 2, 2019.

[29] BCI Competition III: Data set I,

http://www.bbci.de/competition/iii/#data_set_i

[30] Lal, T., T. Hinterberger, G. Widman, M. Schröder, J. Hill, W. Rosenstiel, C. Elger, B. Schölkopf, N. Birbaumer. Methods towards invasive human brain computer interfaces. Proc. of the 17th International Conference on Neural Information Processing Systems (NIPS'04), MIT Press, Cambridge, MA, USA, 2004, pp. 737–744.

***MSc. Eng. Ivaylo Ivaylov**, PhD student, Department. Computer Systems, Faculty of Computer Systems and Technologies.*

e-mail: ivivajlov@tu-sofia.bg

Received on: 29.03.2021

Impact of an energy conservation measure on reducing CO_2 emissions

Mila Ilieva-Obretenova

The study uses saved energy (the expected annual economy of electrical energy) to describe the impact of energy conservation measure (M1) - Reconstruction of the electric drive and control of a basic hoisting machine, on reducing CO_2 emissions. A technological analysis for determining losses without implementation of M1 was made. An investment analysis to determine energy savings in kilowatt-hours for a year with implementation of M1 was performed. The value of energy savings is adjusted with a coefficient of emission factor (CO_2) for electrical energy. Further research will involve an analysis of other energy conservation measures and their impact on the reduction of harmful emissions. In the long run and on a larger scale the improvement of the energy balance of technological systems will be favourable for slowing down climatic changes.

Keywords – CO_2 emissions, coefficient of emission factor, energy conservation measure, saving electrical energy.

Влияние на енергоспестяваща мярка върху намаляване на емисиите на CO_2 (Мила Илиева-Обретенова). *В статията е използвана спестената енергия (очаквана годишна икономия на електрическа енергия) за описание на влиянието на енергоспестяваща мярка (M1) - преустройство на електрозадвижването и управлението на основна подемна машина върху намаляване на емисиите на CO_2. Направен е технологичен анализ за определяне на загуби без внедряване на M1. Извършен е инвестиционен анализ за определяне на енергоспестяването в киловатчаса за една година с внедряване на M1. Стойността на енергоспестяването се коригира с коефициент на емисионен фактор (CO_2) за електрическа енергия. Бъдещата работа включва анализ и на други енергоспестяващи мерки и тяхното влияние върху намаляването на вредни емисии. Дългосрочно и в по-широк мащаб усъвършенстването на енергийния баланс на технологичните системи ще благоприятства забавянето на климатичните промени.*

Introduction

Energy conservation is the effort made to reduce the consumption of energy by using less of an energy service. This can be achieved either by using energy more efficiently (using less energy for a constant service) or by reducing the amount of service used (for example, by driving less). Energy conservation is a part of the concept of Eco-sufficiency. Energy conservation measures (ECMs) reduce the need for energy services and can result in increased environmental quality, national security, personal financial security, and higher savings [11]. Energy saving is achieved through various energy conservation measures which are at the top of the sustainable energy hierarchy [12].

The literature looks at specific factors which negatively affect the environment when certain technological requirements, such as static electricity, are not observed [7]-[9], [15]-[18], [23], [28]-[34], its impact on various processes [3], [4] and devices for neutralizing its harmful effects [10], but no investment analysis of energy conservation measures has been made. Some studies [1], [2], [24]-[27], [37], [38], [39] analyse operations in an enterprise in the mining industry, which implements energy conservation technologies to reduce energy consumption of production, improve processes in energy management and thus favourably impact sustainable ecological development and reduce negative environmental impact but no investment analysis was presented in full. Other studies [5], [6],

[19]-[22], [35], [36] examine the harmful impact of the asymmetry of voltages and currents in three-phase electric circuits but do not make an investment analysis of possible energy conservation measures. The lack of an investment analysis demonstrates that the positive effect of a particular energy conservation measure on the environment cannot be specified quantitatively.

The aim of the present study is to investigate the impact of an energy conservation measure on reducing harmful CO_2 emissions. The measure that was chosen is replacement of the electric drive of a hoisting machine with a new electric drive using the frequency converter - asynchronous motor system, based on the property of this type of motors to be regulated by changing the frequency of their supply voltage. That measure duly demonstrates the improvement of the energy balance of the system. The study may be of interest for university lecturers, university students, researchers, and managers in industry.

Method

All calculations concerning the energy conservation measure, which include the technical indicators of the selected materials, installations/systems and equipment are based on data from research and implementation of a method for calculating savings of electrical energy [13] in accordance with the regulatory framework in the Republic of Bulgaria [14], which has been synchronized with the regulatory framework of the EU regarding energy efficiency.

The economic analysis is based on the prognosticated values of the investment. Prognosticated summed prices of the electrical energy are used to calculate the main economic and financial indicators of the studied measure. Methodology for determining the energy conservation potential of investment is used to calculate the average percentage of energy saving. The reduction of greenhouse gas emissions is calculated based on energy saved for one year [13].

Results

The results of the study comprise technological analysis of a mine hoisting system, an investment analysis for implementation of an energy conservation measure and calculation of the reduction of greenhouse gas emissions.

I. Technological analysis of mine hoisting equipment

Energy indicators of electric drive

The studied hoisting system is driven by an asynchronous motor with rheostat control in the phase (wound) rotor and the speed regulation is performed with great losses of energy (dissipated as heat) in resistors. These losses are so big in accelerating or slowing down the movement of the conveyance that they can become bigger than the needed useful power. For that reason, the energy efficiency of that drive depends on the depth of the hoisting and the number of hoisting cycles per unit of time.

Drive system with frequency regulator

Electric drive using the frequency converter - asynchronous motor system is based on the property of this type of motors to be regulated by changing the frequency of their supply voltage. That voltage is generated by a special electronic converter.

Modern frequency converters have an extremely high coefficient of efficiency. The functional schematic of an example drive system with frequency control of asynchronous motor, which can replace the existing system of asynchronous motor with rheostat control in the phase rotor (Fig.1a), is illustrated in Fig.1b. The two systems have a fundamental difference in regulating the frequency of rotation. Since that frequency of rotation is a component of the output mechanical power together with the moment of rotation, it is of great importance for the energy balance of the hoisting. In the existing regulating system, the energy coming from the network is with a maximum speed and the regulation is done by redirecting part of it to the rotor rheostats. Unlike it the frequency converter - asynchronous motor system receives a regulated amount of energy corresponding to the respective frequency of rotation.

Comparative assessment of the electrical energy consumed when changes in the drive are made

The analysis of the consumed electrical energy in one hoisting cycle is based on a three-period (trapezoidal) diagram of the speed (tachogram) - Fig.2.

Three time intervals are set out in this diagram – of acceleration, of uniform motion and of slowing down as well as of maximum speed (Fig. 2a). Their values are as follows: $t_1=13s$; $t_2=44s$; $t_3=27s$; $V_m=5.48m/s$.

Given these values, accelerations a_1 and a_3, corresponding to the increase of speed to its maximum value and its drop to zero, are calculated:

(1) $$a_1 = V_m/t_1 = 0.422\ m/s^2;$$

(2) $$a_3 = V_m/t_2 = 0.203\ m/s^2.$$

The more complex real tachograms of skip hoists are converted to a three-period diagram.

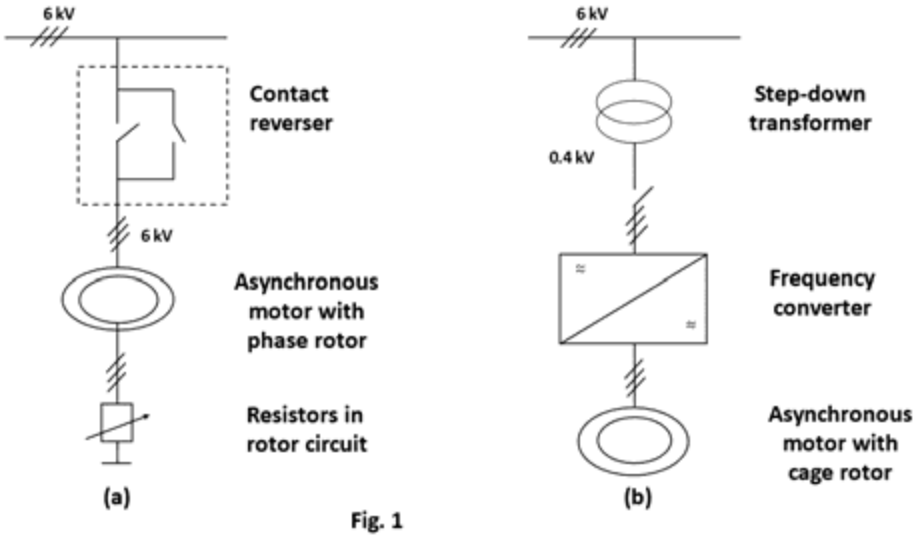


Fig.1. 1a: Asynchronous motor with rheostat driving in phase rotor; 1b: Asynchronous motor with frequency driving.

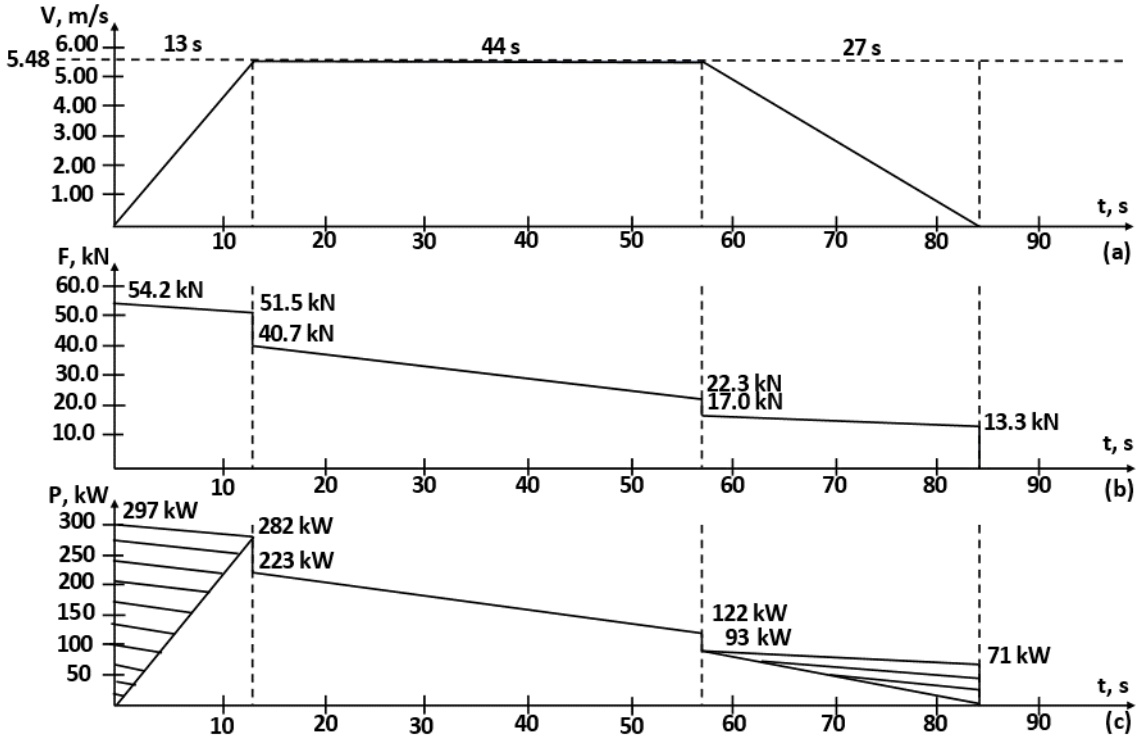


Fig.2. Time diagrams: 2a: Time diagram of speed; 2b: Time diagram of force; 2c: Time diagram of power.

Determining forces in motion

The force during the motion of the conveyance is the sum total of the static force and the dynamic force:

$$F = F_s + F_d, \ N. \tag{3}$$

The static force is determined from the following expression:

$$F_s = \left[kQ + p\left(H - 2x\right)\right]g, \ N, \tag{4}$$

where: k – coefficient accounting for negative resistance caused by the bending of the rope, friction in the bearings and the guides and air resistance in the motion of the skip. As a generally accepted value is taken k = 1.2; Q – mass of the mineral in the skip,

Q = 2630 kg; p – weight of 1 meter of rope, p = 3.89 kg/m; H – depth of hoisting, H = 326 m; x – coordinates of the skip; g – acceleration due to gravity.

Dynamic force occurs only during periods of acceleration and is determined by the expression:

(5) $$F_d = (m_1 + m_2)a, \ N,$$

where: m_1 – the overall mass of all parts in translational motion, m_1 = 8484 kg (that mass is the sum total of the mass of the minerals, the two skips and the ropes); m_2 – inertial mass of all rotating parts reduced to the circumference of the drum, m_2 = 17300 kg (taken from the catalogue of the hoisting machine).

Determining the force at the beginning of the motion

The time diagram of the forces calculated below is given in Fig.2b.

(6) $$F_1' = F_{s1}' + F_{d1}, \ N,$$

where: F_{s1}' - static force at the beginning of the motion,

(7) $$F_{s1}' = (kQ + pH)g = 43400 \ N;$$

F_{d1} - dynamic force in accelerating the conveyance to reach its maximum speed,

(8) $$F_{d1} = (m_1 + m_2)a_1 = 10873 \ N;$$

(9) $$F_1' = 54273 \ N.$$

Determining the force at the end of the accelerated motion

With $x = h_1 = 0.5V_m t_1$ = 35.6 m

(10) $$F_1'' = F_{s1}'' + F_{d1} \ N,$$

(11) $$F_{s1}'' = [kQ + p(H - 2h_1)]g = 40684 \ N,$$

(12) $$F_1'' = 51557 \ N.$$

Determining the force at the beginning of the uniform motion with maximum speed

Here the dynamic force becomes zero due to the uniform motion and the force is equal only to the static force at the end of the acceleration motion:

(13) $$F_2' = F_{s1}'' = 40684 \ N.$$

Determining the force at the end of the uniform motion with maximum speed

At this point there is only a static force, but it is smaller than the one at the beginning of the uniform motion due to the winding of the rope. Thus, the force is determined by the difference between the force at the beginning of the uniform motion and the force determined by the wound rope:

(14) $$F_2'' = F_{s1}'' - 2gV_m t_2 p = 22280 \ N.$$

Determining the force at the beginning of the period of braking (retarded motion)

That force is determined by subtracting the dynamic force for a period of delay from the force at the end of the uniform motion with maximum speed:

(15) $F_3' = F_2'' - F_{d3} \ N$, F_{d3} -

dynamic force for a period of delay,

(16) $$F_{d3} = (m_1 + m_2)a_3 = 5236 \ N,$$

(17) $$F_3' = 17044 \ N.$$

Determining the force at the end of the period of braking (retarded motion)

With $x = H = 326 \ m$,

(18) $$F_3'' = F_{s3}'' - F_{d3} \ N,$$

(19) $$F_{s3}'' = (kQ - pH)g = 18520 \ N,$$

(20) $$F_3'' = 13284 \ N.$$

Diagram of the powers

The time diagram of the powers is given in Fig.2c

and corresponds to values calculated at this point. The power of the shaft of the drum is determined as the product of the multiplication of the force and speed for the respective point:

(21) $P = FV.10^{-3}\ kW$.

Thus, we get the following characteristic points in the diagram of the power: $P_1^{'} = 0\ kW$,

(22) $P_1^{``}=F_1^{``}V_m.10^{-3}=282.5kW,$

(23) $P_2^{'} = F_2^{'}V_m.10^{-3} = 223\ kW$,

(24) $P_2^{"} = F_2^{"}V_m.10^{-3} = 122\ kW$,

(25) $P_3^{'} = F_3^{'}V_m.10^{-3} = 93\ kW$, $P_3^{"} = 0\ kW$.

For hoisting driven through the asynchronous motor – rotor rheostat system the diagram of the consumed electrical power is similar to the diagram of the forces and is determined according to the expression:

(26) $$P = \frac{FV_m}{\eta_{red}\eta_{dv}}.10^{-3}\ kW .$$

The denominator comprises the coefficient of efficiency of the reducer and the engine. In order to simplify the expression, the coefficient of efficiency is not considered as the coefficient of efficiency of the reducter and the engine in the old and the new drive system are the same. Thus, we get the following characteristic points in the diagram of the consumed electrical power:

(27) $P_1^{'} = F_1^{'}V_m.10^{-3} = 297\ kW$,

(28) $P_1^{``} = F_1^{``}V_m.10^{-3} = 282.5\ kW,$

(29) $P_2^{'} = F_2^{'}V_m.10^{-3} = 223\ kW$,

(30) $P_2^{"} = F_2^{"}V_m.10^{-3} = 122\ kW$,

(31) $P_3^{'} = F_3^{'}V_m.10^{-3} = 93\ kW$,

(32) $P_3^{"} = F_3^{"}V_m.10^{-3} = 71\ kW$.

From the power diagram we can find the total energy for one hoisting cycle. This value corresponds to the area under the graph of the engine power. It is expressed as the sum total of the areas of the three trapezes:

(33) $W_e = 13567\ kW.s$

The losses in the resistors are determined as the difference between the total energy W_e and the energy at the shaft of the drum. These losses correspond to the sum total of the areas of the two triangles hatched with lines:

(34) $W_z = 2889\ kW.s$.

The losses in the rotor resistors are calculated as follows:

(35) $W_z.100/W_e = 21.3\ \%.$

The analysis demonstrates that with the implementation of a frequency control drive, where there are no rotor resistors and respectively no losses in them, this drive will be more economical by more than 20 %.

Factors, which improve the energy efficiency of the hoisting: stepless speed regulation; considerable reduction of losses in the control panels; lower inertia moment for the engine rotor; increased reliability of the drive system; high efficiency of the hoisting machine in maintenance checks.

Factors, which decrease energy efficiency of hoisting: availability of converters of electric energy before it enters the engine, frequency converter and in some schemes a transformer.

These factors to a certain extent compensate for one another and would not significantly change the above conclusion about energy efficiency.

II. An investment analysis

The analysis of the operations on site reveals the use of old, ineffective equipment with high energy consumption – shaft of the "Kapitalna" type. It was produced in 1961 according to standard GOST3006-52. That hoisting equipment was designed in the 1950s. It is driven by an asynchronous motor with rheostat control in the phase (wound) rotor. At the time of its design energy efficiency (the rational use of electrical energy) and the coefficient of efficiency were not issues of priority.

The minimum requirements for M1 are:

- Electric drives ≤ 375 kW: Standard of energy efficiency IE2 or IE3.

Description of state

The calculations of the investment analysis, for clear representation, are given in a tabular form.

The baseline is presented in Table 1.

Proposed measure M1

The measure involves activities related to the replacement of the electric drive of the hoisting equipment in "Kapitalna" shaft with new electric drive using the frequency converter - asynchronous motor system, based on the property of this type of motors to be regulated by changing the frequency of their supply voltage.

The technical analysis includes calculations of energy savings based on the modes of operation and the installed capacity of the current and the prognosticated state after the introduction of the energy conservation measure – Table 2.

Capital expenditure and operation and maintenance costs

The economic analysis of the measure is done based on preliminary estimates of capital expenditure, prognosticated summed prices of electrical energy and zero costs for maintenance and operation.

The calculations of the financial analysis are presented in Table 3.

Table 1

Baseline

	Description	**Value**	**Dimension**	**Explanation**	**Notes**
A	Installed capacity – "Kapitalna" hoist shaft	320	kW		technical data
B	Annual operation	1043	h/y		energy analysis
C	Annual productivity	110741	t/y		energy analysis
D	Annual consumption of electrical energy	333 753	kWh/y	= A x B	Calculated
E	Specific annual consumption of electrical energy	3.0	kWh/t	= D / C	Calculated

Table 2

After implementation of M1

	Description	**Value**	**Dimension**	**Explanation**	**Notes**
F	Operation capacity - "Kapitalna" hoist shaft	256	kW		Calculated
G	Annual operation	1043	h/y		energy analysis
H	Annual productivity	110741	t/y	= C	energy analysis
I	Annual consumption of electrical energy	267 002	kWh/y	= F x G	calculated according to methodology
J	Specific annual consumption of electrical energy	2.4	kWh/t	= I / H	Calculated

Table 3

Basic indicators in the financial analysis of M1

	Economies	Value	Dimension	Explanation	Notes
K	Economy of electrical energy	66 751	kWh/y	= D – I	Calculated
L	Specific annual economies	0.60	kWh/t	= E – J	Calculated
M	Average price of electrical energy	0.120	BGN/kWh		Expert estimate
N	Annual economy of electrical energy	8 010	BGN/y	= K x M	Calculated
O	Annual economy of economic activity	51664.13	BGN/y		Calculated
P	Overall annual economy	59 674.13	BGN/y	= N + O	Calculated
Q	Investment for ECM	344 427.53	BGN		Investment intention
R	Redemption period	5.8	y (year)	= Q / P	Calculated
S	Energy saving	20.00	%	= K x 100/ D	Calculated

Since the object of study is an enterprise the annual saving of energy from the implemented measure is referred to the production of one unit. All prices do not include VAT. After the implementation of the proposed measure consumption was reduced by 66751 kWh/y.

Specific consumption and energy savings

Specific power consumption and determination of energy savings:

- Specific power consumption (current state): 3.0 kWh/t;
- Specific power consumption (after the implementation of M1): 2.4 kWh/t;
- Energy savings from the implemented measure M1:

(36) ((3 – 2.4) / 3) x 100 = 20 %.

III. Calculation of the reduction in greenhouse gas emissions

The estimate was made by adjusting the saved energy with the coefficient of emission factor for electrical energy

(37) $K_{el} = 1.039\ kgCO_2 / kWh.$

The result for expected annual economy of electrical energy from $W_S = 66751\ kWh / y$ in saved emissions S_{CO_2} is 69.4 tonnes of CO_2 per year:

(38) $S_{CO_2} = W_S . K$

(39) $S_{CO_2} = W_S . K_{el} . 10^{-3} = 69{,}4\ tCO_2 / y.$

Conclusion

The energy conservation measure M1 – Reconstruction of the electric drive and control of a main hoisting device in a mining enterprise results in reduction of annual emissions of CO_2. Further research will involve analysis of other energy conservation measures and their impact on the reduction of harmful emissions. In the long term and on a larger scale the improvement of the energy balance of the technological systems will be favourable for slowing down climatic changes.

REFERENCES

[1] Prodanov, I., K. Trichkov, A. Zabchev, R. Aleksandrov. Approaches for implementation of energy-saving technologies in the mining industry. Mining work and Geology Journal, Sofia, 1-2 / 2014, pp. 53-56. (in Bulgarian)

[2] Zabchev, A., R. Aleksandrov, I. Prodanov. Analysis of energy efficiency in the reconstruction of the electric drive of a lifting machine. Proceedings of the National scientific and technical conference with international participation Automation in Mining Industry and Metallurgy, BULCAMC`14, Sofia, 06 - 07 November 2014, pp. 122-127. (in Bulgarian)

[3] Prodanov, I. Study of the electrisation of transformer lubricating oil. Mining Work and Geology Journal, Sofia, 3-4 / 2014, pp. 53-56. (in Bulgarian)

[4] Prodanov, I. The Problems of Static Electricity during Printed Circuit Board Assembly. Proceedings of XV International Conference on Lighting BulLight 2014, Sozopol, 10 - 13 June 2014, pp. 197-201.

[5] Stefanov, S., I. Prodanov. Asymmetry of Voltages and Currents in Three Phase Electric Circuits and Methods of Reducing Them. Proceedings of XV International Conference on Lighting BulLight 2014, Sozopol, 10 - 13 June 2014, pp. 202-205.

[6] Stefanov, S., I. Prodanov. Electric energy and quality. Asymmetric three phase circuits. Electro safety. Monography. Publishing Avangard Prima, Sofia, 2016. (in Bulgarian)

[7] Trichkov, K., I. Prodanov. Guide for laboratory exercises on electro materials, Publishing house St. Ivan Rilski, Sofia, 2013. (in Bulgarian)

[8] Stefanov, S., I. Prodanov. Static electricity – theory and praxis, Monography. Publisher Avangard Prima, Sofia, 2013. (in Bulgarian)

[9] Stefanov, S., I. Prodanov. Device for Measuring the Potential, Intensity and Charge of the Field of Electrostatically Charged Materials. Journal of the Technical University – Sofia, Plovdiv branch, Fundamental Sciences and Applications, Vol. 19, B. 1, International Conference Engineering, Technologies, and Systems TECHSYS, 2013, pp. 227-230.

[10] Prodanov, I. Research on electrostatic field parameters and development of electromechanical devices for neutralisation of its harmful impact on technological processes, PhD Thesis, Publishing House St. Ivan Rilski, Sofia, 2006. (in Bulgarian)

[11] Zehner, Ozzie. Unintended Consequences. In Green Technology, edited by Paul Robbins, Dustin Mulvaney and J. Geoffrey Golson, 427-32. London: Sage, 2011.

[12] A guide for effective energy saving. Renewable Energy World. Retrieved 14 June 2016.

[13] EN EM KO Ltd. www.enemko.com. Reports of exploration for energy efficiency of industrial systems, 2005 – 2020. (in Bulgarian)

[14] Energy Efficiency Act, [Online]. Available: https://www.me.government.bg/en/library/zakon-za-energiinata-efektivnost-537-c25-m1515-2.html

[15] Stefanov, S., I. Prodanov. Measurement and Calculation of the Electrifying Current of Electrified Oil Products Using the Sampling Method. University of Mining and Geology St. Ivan Rilski, Journal of Mining and Geological Sciences, Vol. 61, Part III Mechanization, Electrification and Automation in Mines, Sofia, 2018, pp. 47-49.

[16] Prodanov, I., S. Stefanov. Measurement and Calculation of the Density of the Charge of Electrified Oil Products Using the Sampling Method. University of Mining and Geology St. Ivan Rilski, Journal of Mining and Geological Sciences, Vol. 61, Part III Mechanization, Electrification and Automation in Mines, Sofia, 2018, pp. 44-46.

[17] Stefanov, S., I. Prodanov. Charge Relaxation in a Reservoir Filled with Electrified Liquid. University of Mining and Geology St. Ivan Rilski, Journal of Mining and Geological Science, Part III Mechanization, Electrification and Automation in Mines, Vol. 60, 2017, pp. 52-54.

[18] Stefanov, S., I. Prodanov. Charge Accumulation in the Process of Filling of Electrified Liquid inside a Reservoir. University of Mining and Geology St. Ivan Rilski, Journal of Mining and Geological Science, Part III Mechanization, Electrification and Automation in Mines, Vol. 60, Sofia, 2017, pp. 49-51.

[19] Stefanov, S., I. Prodanov. Analytical Methodology for Calculating Asymmetry of Currents in Three-Phase Electrical Systems. Annual Proceedings of the University of Mining and Geology St. Ivan Rilski, Vol. 57, Part III Mechanization, electrification and automation in mines, Sofia, 2014, pp. 51-53.

[20] Stefanov, S., I. Prodanov. Equivalent Scheme of a Three-Phase Circuit with Insolated Neutral, When Calculating the Current Flowing Through the Human Body at the Touch of One of the Phases, at Large Conductivity Rates of the Insulation. Annual Proceedings of the University of Mining and Geology St. Ivan Rilski, Vol. 57, Part III Mechanization, electrification and automation in mines, Sofia, 2014, pp. 47-50.

[21] Stefanov, S., I. Prodanov. Allocation of the Currents with Zero Sequence along the Cable Line in Single-Phase Ground Faults. Annual Proceedings of the University of Mining and Geology St. Ivan Rilski, Vol. 56, Part III Mechanization, electrification and automation in mines, Sofia, 2013, pp. 82-85.

[22] Stefanov, S., I. Prodanov. Signal Graphs and Their Implementation for Resolving Issues Regarding Electrical Safety in Three-Phase Electrical Circuits. Annual Proceedings of the University of Mining and Geology St. Ivan Rilski, Vol. 56, Part III Mechanization, electrification and automation in mines, Sofia, 2013, pp. 78-81.

[23] Stefanov, S., I. Prodanov. Study of Explosion-Proof Properties of Inductive Neutralizers. Journal of the Technical University – Sofia, Plovdiv branch, Fundamental Sciences and Applications , Vol. 19, B. 2, International Conference Engineering, Technologies, and System TECHSYS, 2013, pp. 285-288.

[24] Minekov, N., R. Istalianov, I. Jochev, I. Prodanov, I. Mladenova, N. Lakov. Determination of the compensatory ability of synchronous engines taking into consideration the influence of the voltage on the electrical power supply network. Annual Proceedings of University of Mining and Geology St. Ivan Rilski, Vol. 53, Part III Mechanization, electrification and automation in mines, Sofia, 2010, pp. 108-110.

[25] Kartselin, E., I. Mladenova, R. Istalianov, I. Prodanov, I. Jochev, N. Minekov. Influence of the high order harmonics on the behavior of the Three-Phase machine. Annual Proceedings of University of Mining and Geology St. Ivan Rilski, Vol. 53, Part III Mechanization, electrification and automation in mines, Sofia, 2010, pp. 105-107.

[26] Istalianov, R., I. Jochev, I. Prodanov, N. Minekov, I. Mladenova. Experimental study on the main water pipe system intended for deriving water/main pump installation, located within underground mine named RudmetalCorporation. Annual Proceedings of University of Mining and Geology St. Ivan Rilski, Vol. 53, Part III Mechanization, electrification and automation in mines, Sofia, 2010, pp. 84-86.

[27] Kartselin, E., R. Istalianov, I. Prodanov, I. Jochev, N. Minekov, I. Mladenova. Use of main pump installations in underground mines as users-regulators of power. Annual Proceedings of University of Mining and Geology St. Ivan Rilski, Vol. 53, Part III Mechanization, electrification and automation in mines, Sofia, 2010, pp. 77-80.

[28] Stefanov, S., I. Milev, I. Prodanov. Use of the Shockly-Ramo theorem for determination of the string neutralizer charge per unit of length. Annual Proceedings of University of Mining and Geology St. Ivan Rilski, Vol. 48, Part III Mechanization, electrification and automation in mines, Sofia, 2005, pp. 89-93.

[29] Stefanov, S., I. Milev, I. Prodanov. Inductive neutralizer of bulk charges with increased efficiency. Annual Proceedings of University of Mining and Geology St. Ivan Rilski, Vol. 48, Part III Mechanization, electrification and automation in mines, Sofia, 2005, pp. 89-93.

[30] Stefanov, S., I. Milev, I. Prodanov. Experimental Research in Explosion-Proof Inductive Neutralizer with Discharge Element a Mobile Conductive Sphere. Constantin Brancusi University of Targu Jiu - Engineerig Faculty Scientific Conference – 9th edition - with international participation, November 5 - 6, 2004 - Targu Jiu, Romania, pp. 491-496.

[31] Stefanov, S., I. Milev, I. Prodanov. Explosion-Proof Inductive Neutralizer with Discharge Element a Mobile Conductive Sphere. Constantin Brancusi University of Targu Jiu – Engineering Faculty Scientific Conference - 9 edition - with international participation, November 5 - 6, 2004 - Targu Jiu, Romania, pp. 497-503.

[32] Stefanov, S., I. Milev, I. Prodanov. Features of the insulating liquids electrification from overflowing main pipelines to receiving tanks and devices for the appeared static electricity elimination. Annual Proceedings of University of Mining and Geology St. Ivan Rilski, Vol. 47, Part III Mechanization, electrification and automation in mines, Sofia, 2004, pp. 73-75.

[33] Stefanov, S., I. Milev, I. Prodanov. Measurement by a non-contact method of the electrostatic field parameters of flat dielectrics. Annual Proceedings of University of Mining and Geology St. Ivan Rilski, Vol. 47, Part III Mechanization, electrification and automation in mines, Sofia, 2004, pp. 69-72.

[34] Stefanov, S., I. Milev, I. Prodanov. Inductive neutralizer of static electricity with double capacitive coupling. Mining work and Geology Journal, Sofia, 6 / 2004, pp. 27-30. (in Bulgarian)

[35] Stefanov, S., I. Milev, I. Prodanov. Electrical energy and quality, Energy and business Journal, Sofia, XII 2003, pp. 30-32. (in Bulgarian)

[36] Stefanov, S., T. Hristova, I. Prodanov. Guide for laboratory exercises on theoretical electrotechnics, Publisher Avangard Prima, Sofia, 2013. (in Bulgarian)

[37] Jack Jeswiet, Alex Szekeres. Energy Consumption in Mining Comminution. Procedia CIRP, Vol. 48, 2016, pp. 140-145, ISSN 2212-8271, https://doi.org/10.1016/j.procir.2016.03.250.

[38] Bogunovic, D., V. Kecojevic, V. Lund, M. Heger, P. Mongeon. Analysis of Energy Consumption in Surface Coal Mining. Transactions of the Society for Mining, Metallurgy, and Exploration, Inc. 326, 2009, pp.79-87.

[39] Kecojevic, V., I. Vukotic, D. Komljenovic. Production, Consumption and Cost of Energy for Surface Mining of Bituminous Coal. Mining Engineering, 66, 2014, pp. 51-57.

Assoc. Prof. PhD. Mila Ilieva-Obretenova *– University of Mining and Geology "St. Ivan Rilski"- Sofia, Department of Electrical Engineering, Electronics Laboratory, scientific interests: Industrial Electronics, Semiconductor Devices and Integrated Circuits, Nanoelectronics.*

tel.: 0889 41 81 75 e-mail: mila.ilieva@mgu.bg

Received on: 14.04.2021

♦ SECOND CALL FOR PAPERS

The ELMA 2021 organizers invite you to submit papers in the area of the conference topics. A two-pages extended abstract for each paper should be submitted via the conference management system. Shorter abstracts will not be processed. Abstracts will be reviewed by at least two reviewers. Each author can participate in up to three papers.

♦ SCOPE OF THE CONFERENCE

The first conference ELMA was held in Varna in November 1975. Originally the conference was conceived as a forum on electric machines. Gradually, the scope is extended and currently it includes all aspects of modern electrical engineering.
The main aim of the conference ELMA 2021 is to give an opportunity to academics, scientists, engineers, PhD students, manufacturers and users to discuss and exchange information about the results of their theoretical and experimental investigation, the utilization of new materials and application of new technologies in the field of electrical machines, drives, power electronics and power systems.

♦ VENUE

The 17th International Conference on Electrical Machines, Drives and Power Systems ELMA 2021 will be held in Sofia, the Capital City of Bulgaria, in the Technical University, Faculty of Electrical Engineering, Bl. 12.

♦ LANGUAGE

The working language of the conference is English.

♦ PAPER PRESENTATION

Papers will be presented in oral and poster sessions. Authors are asked to suggest the preferred form of presentation while the Program Committee will make the final decision.
Authors are expected to attend the conference in person to present their papers and share their ideas. Depending on the pandemic situation, the conference may be held in attendance, either online or mixed - in attendance and online.

♦ IEEE PUBLICATION

IEEE is the copyright holder of the Conference Proceedings. All accepted papers will be published in the Conference Proceedings and submitted to the IEEE *Xplore*® Digital Library, Scopus and Web of sciences databases. IEEE reserves the right to exclude a paper from distribution after the conference, including IEEE *Xplore*® Digital Library, if the paper is not presented by the author at the conference.

♦ TOPICS

- Rotating electrical machines and transformers
- Electrical drives and actuators
- Electric traction and electric vehicles
- Electrohydraulic power and control systems
- Control systems in electrical engineering
- Power electronic converters in electrical engineering
- Power systems
- Renewable energy sources and systems
- Energy efficiency and power quality
- Electrotechnologies and electrical appliances
- Modeling and virtual engineering in electrical industry
- ICTs, measurement in electrical engineering
- New materials and components
- Management and marketing in electrical engineering and power industry
- Standardization
- Education, teaching, multimedia, distance and e-learning in electrical engineering

Further information can be obtained on the Conference Web Page: http://elmaconf.tu-sofia.bg

♦ ACCOMPANYING EVENTS

Companies are kindly invited to take part in the conference with presentations, innovations and sample demonstrations. To ensure participation, together with the application, please inform us about: the title of technical performance, the requested area and necessary voltage.

♦ DEADLINES

Extended abstract submission:	**April 16, 2021**
Notification of acceptance:	**May 17, 2021**
Early registration fee deadline:	**May 30, 2021**
Receipt of Full Papers:	**June 14, 2021**

2021 17-th INTERNATIONAL CONFERENCE ON ELECTRICAL MACHINES, DRIVES

AND POWER SYSTEMS

July 1-4 2021
SOFIA, BULGARIA

Organized by:

Union of Electronics, Electrical Engineering and Telecommunications (CEEC)
IEEE Bulgaria Section

with the support of
Technical University of Sofia
Technical University of Varna
Technical University of Gabrovo
University of Ruse
Federation of Scientific and Technical Unions

♦ CONFERENCE CHAIRMEN

Vladimir LAZAROV Technical University of Sofia
Bohos APRAHAMIAN Technical University of Varna

♦ INTERNATIONAL PROGRAM COMMITTEE

Slavoljub ALEKSIC, Serbia
Bohos APRAHAMIAN, Bulgaria
Kiril ALEXIEV, Bulgaria
Yvan AVENAS, France
Viktor BESPALOV, Russia
Gantcho BOJILOV, Bulgaria
Hartmut BRAUER, Germany
Nikolai CHRISTOV, France
Violeta CIUCUR, Romania
Moumen DARCHERIF, France
Dimitar DIMITROV, Bulgaria
Viktor ELISTRATOV, Russia
Masato ENOKIZONO, Japan
Bulent ERTAN, Turkey
Bruno FRANCOIS, France
Željko HEDERIĆ, Croatia
Antonios KLADAS, Greece
Pavel KURBATOV, Russia
Vladimir LAZAROV, Bulgaria
Afef LEBOUC, France
Jiři LETTL, Czech Republic
Iliana MARINOVA, Bulgaria
Nikola MIHAILOV, Bulgaria
Gilles NOTTON, France
Cornel PANAIT, Romania
Ioan POPA, Romania
Emil RATZ, Bulgaria
Robert RATZ, USA
Christian SCHAEFFER, France
Brigite SCHMIDT, Germany
Dietmar SCHMIDT, Germany
Wolfgang SCHROPPEL, Germany
Sergey SHIRINSKY, Russia
Emil SOKOLOV, Bulgaria
Mile STANKOVSKI, Republic of North Macedonia
Hugo VALDERRAMA-BLAVI, Spain
Ivan YATCHEV, Bulgaria

♦ TECHNICAL PROGRAM COMMITTEE

CHAIRMAN:
Zahari ZARKOV CEEC, TU-Sofia

MEMBERS:

Kiril ALEXIEV	IEEE Bulgaria Section
Katya ASPARUHOVA	IEEE Bulgaria Section
Stefan PATCHEDJIEV	CEEC
Ludmil STOYANOV	TU-Sofia
Emil RACHEV	TU-Sofia
Valentin MILENOV	TU-Sofia
Ivan BACHEV	TU-Sofia
Vladislav PETROV	TU-Sofia
Maik STREBLAU	TU-Varna
Pencho VLADIMIROV	TU-Gabrovo
Boris EVSTATIEV	University of Ruse

♦ LOCAL ORGANIZING COMMITTEE

CHAIRS:

George TODOROV
Maik STREBLAU
Ludmil STOYANOV

MEMBERS:

Valentin KOLEV	Plamen RIZOV
Nikolay MATANOV	Radoslav SPASOV
Iliana MARINOVA	Raina TZENEVA
Valentin MATEEV	Ivaylo DOLAPCHIEV
Angel TZOLOV	Dimo STOILOV
Dimitar ARNAUDOV	Nikolay HINOV
Boris EVSTATIEV	Georgi SAVOV
Sava KONTROV	Encho POPOV
Mariya MARINOVA	Marin MARINOV
Tatyana DIMOVA	Georgi ZHELEV
Janita SLAVOVA	Ivaylo NEDELCHEV
Mariana TODOROVA	Emilian BEKOV
Violeta BOZHIKOVA	Hristo VALCHANOV
Yoncho KAMENOV	Valentin GYUROV
Svetlana LESIDRENSKA	Nikolay DJAGAROV
Chavdar ALEXANDROV	Dragiya YANULOV

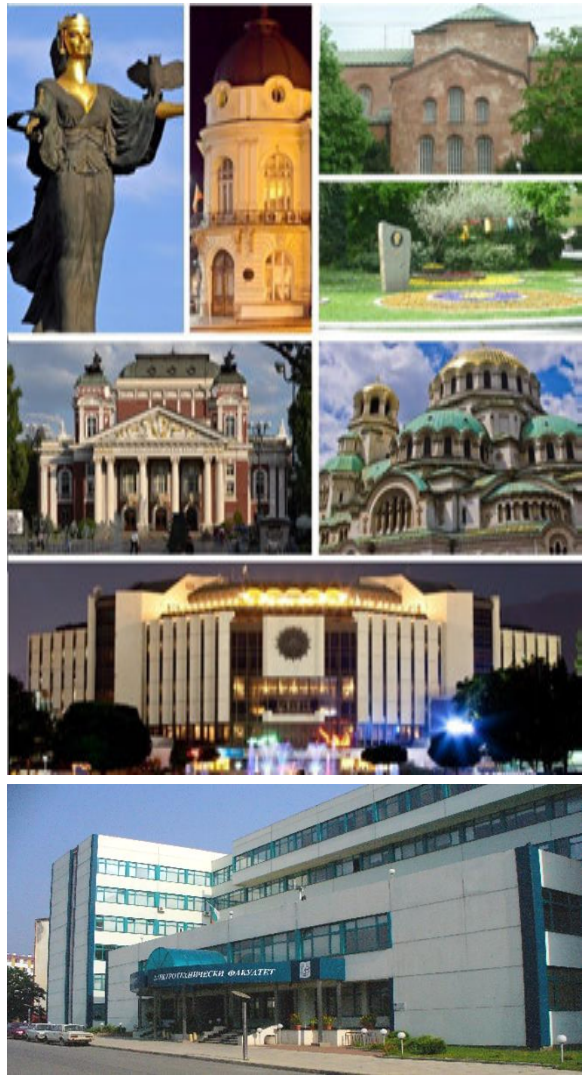
♦ REGISTRATION FEE

Due to the ongoing COVID-19 pandemic crises, the ELMA 2021 conference will be held ONLINE and HYBRID.
The registration fee for all type of participants is **200 EUR**. It is due until 14 June 2021.
The fee covers participation to the online sessions, paper processing, publication in IEEE Xplore database, the welcome cocktail and coffee breaks.

♦ ADDRESS FOR CORRESPONDENCE

Technical University of Sofia
Dept. of Electrical Machines, Bl. 12
For ELMA 2021
8, Kliment Ohridski Blvd.
1000 SOFIA, BULGARIA

Phone: (+359 2) 965 2465, (+359 2) 987 9767
Fax: (+359 2) 987 9360
E-mail: elma2021@tu-sofia.bg
http://elmaconf.tu-sofia.bg



ЕЛЕКТРОТЕХНИКА И ЕЛЕКТРОНИКА E+E

56 год. 3-4/2021

Научно-техническо списание

Издание на: **Съюза по електроника, електротехника и съобщения /CEEC/**

Адрес:
ул. “Раковски” № 108
ет. 5, стая 506
София 1000
тел.: +359 2 987 97 67
e-mail: epluse@ceec.bg
http://epluse.ceec.bg

ISSN 0861-4717

С Ъ Д Ъ Р Ж А Н И Е

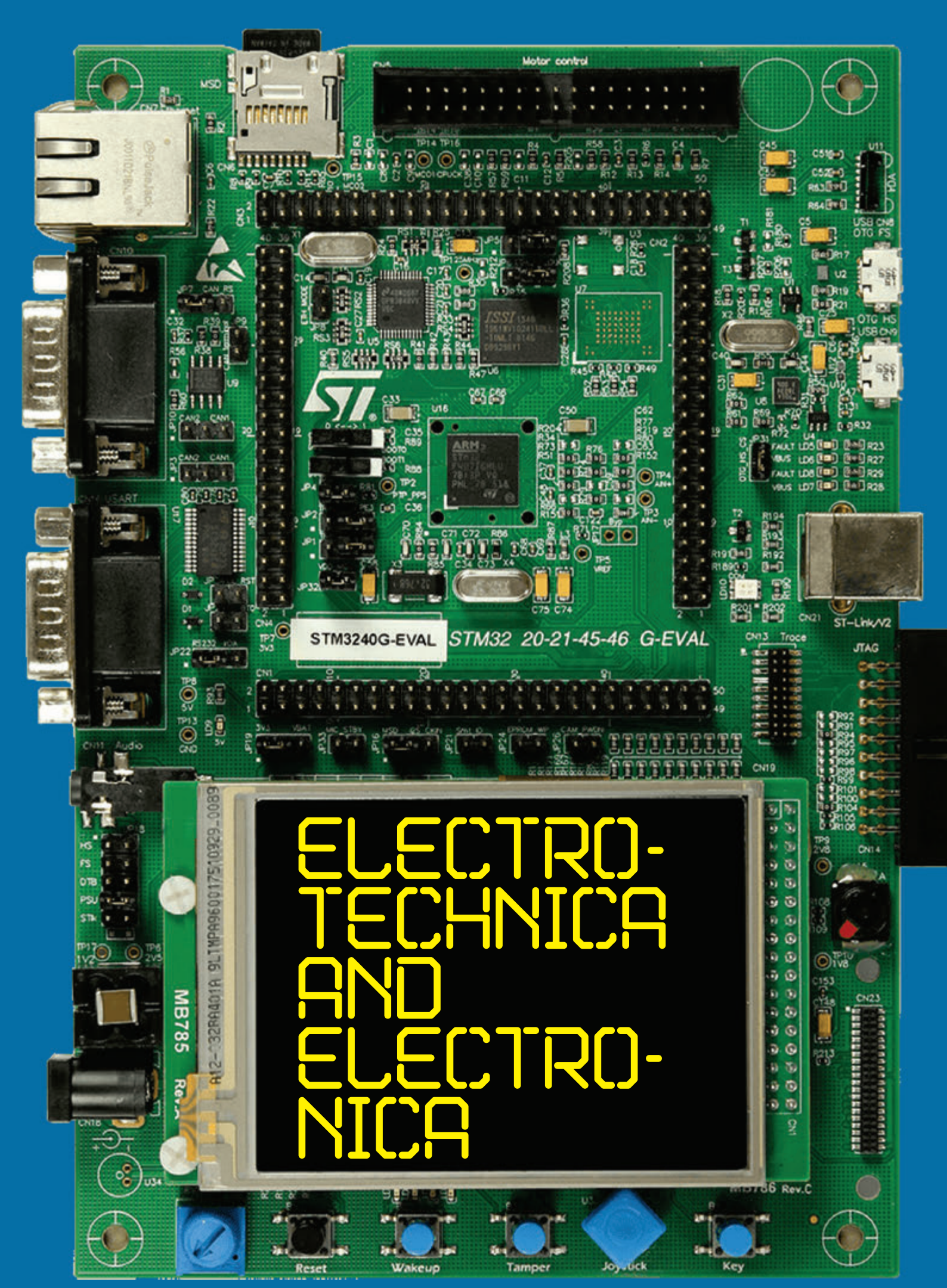
КОМПЮТЪРНИ НАУКИ

Красимира Швертнер
Обществен и частен облак
с автономни бази от данни 29

Ивайло Е. Ивайлов
BCI класификации, използващи генетичен
алгоритъм и k-fold кръстосана валидация 38

ЕЛЕКТРОТЕХНИКА

Мила Илиева-Обретенова
Влияние на енергоспестяваща мярка
върху намаляване на емисиите на CO_2 46



JOURNAL E+E, 108 RAKOVSKI ST, 1000 SOFIA, BULGARIA, EPLUSE@CEEC.BG, WEB: EPLUSE.CEEC.BG