

# A Computing System for Processing the Interstate Power Grid Data

I. Trofimov\*, Sergey V. Podkovalnikov, Leonid N. Trofimov, Ludmila Y. Chudinova

Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences, Irkutsk, Russia

**Abstract** – The paper presents a Data Processing and Geo-information Computing System (DPGICS) intended to study and forecast the expansion of interstate power grids. We propose an original technology for data storage in an object-oriented database, and a technology for data processing and representation using user-friendly interface. An example is given to demonstrate the DPGICS applied to visualize energy data on maps, and to form aggregated tables with the results obtained using the ORIRES optimization model being part of DPGICS.

**Index Terms:** geo-information system, optimization model, data processing, object database, power plant, power system, interstate power interconnection.

## I. INTRODUCTION

Establishment of interstate power grids (IPGs) is a global trend in the world electric power development. This process involves proving their effectiveness, projecting their further expansion, meeting the economic interests, complying with technical standards of the member countries, and considering many other aspects.

Investigations of such kind require big preliminary work on collection and analysis of a huge amount of data (technical, economic, ecological), complex technical and economic calculations, significant intellectual and information resources. Without developed problem-oriented software reflecting the specificity of the given subject area, investigations take a great deal of time and effort. Multi-aspect results of calculations can hardly be perceived by experts without a convenient user interface for data processing and presentation.

Currently, methods of processing and analysis of large

amounts of data are studied by many researchers from different countries in various science areas: computer science, mathematics, power engineering and others. In Russia, the well-known studies in the area of data processing were conducted by A.A. Barsegyan, M.S. Kupriyanov, V.V. Stepanenko and I.I. Holod – "The Data Analysis Technologies. OLAP Data Mining" [1], V.A. Duke – "Telemedicine", "Data Mining" [2]. The foreign reserachers involved in the study of information technologies for intellectual data processing (Data Mining) are Jiwei Han and Philip S. Yu (The Data Mining Group, University of Illinois, USA), and Charu Aggarwal (IBM, USA) to name but a few.

Many researchers and international organizations investigate the Interstate Power Grids. These are Global Energy Interconnection Development and Cooperation Organization (GEIDCO, China), Asia Pacific Energy Research Centre (APEREC), United Nations Economic and Social Commission for Asia and the Pacific (UN ESCAP), Renewable Energy Institute (Japan) [3], Korea Electrotechnology Research Institute (South Korea) [4-5], China Electric Power Planning and Engineering Institute (China), Economy and Technology Research Institute (China), Mongolia Energy Regulatory Commission [6] and others. A lot of universal software was developed. In most cases, relational databases are used for energy data storage. OLAP, Data Mining and others methods are used to process multidimensional data sets. The proposed methods of data storing and processing are intended for universal solutions, but they do not meet the goals set.

Comprehensive research of transmission lines and electricity generation is indispensable for the IPG system expansion and substantiation [7]. Therefore, an optimization model is used to study the IPG expansion, choose the optimal commissioning of power plants and transmission lines to cover the growing load in the target year [8]. We collect and process huge arrays of data for this model. However, most energy databases are not publicly available for research in this domain. We needed software with convenient user interface to work with the optimization model and to analyze the results obtained.

Herewith, the IPG and electric power systems (EPSs)

\* Corresponding author.

E-mail: [t\\_john88@isem.irk.ru](mailto:t_john88@isem.irk.ru)

<http://dx.doi.org/10.25729/esr.2018.03.0008>

Received September 10, 2018. Revised November 28, 2018.

Accepted December 29, 2018. Available online January 25, 2019.

© 2018 ESI SB RAS and authors. All rights reserved.

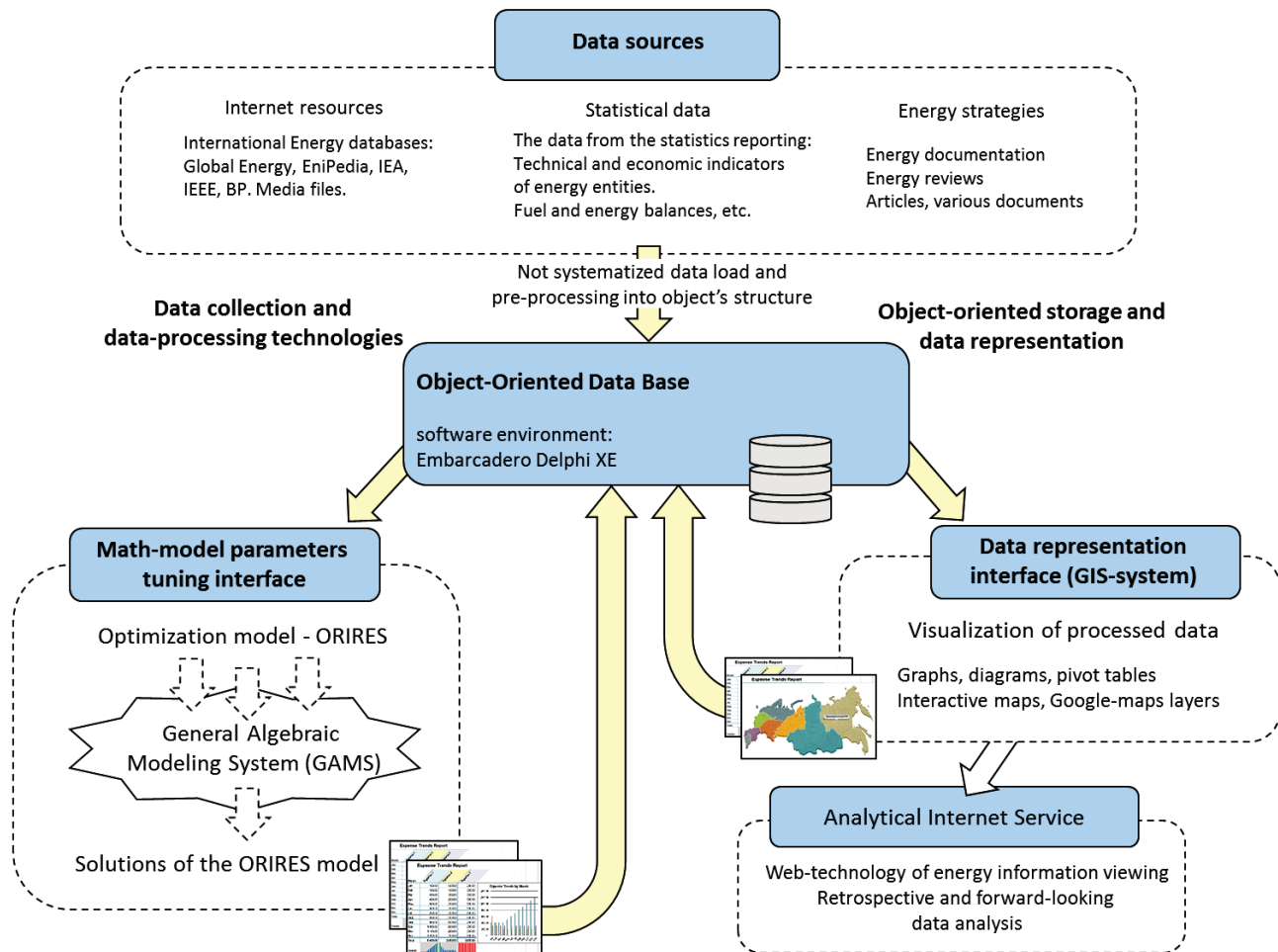


Fig. 1. DPGICS functional blocks.

being their parts are modeled as structurally complex power entities (EPS, power plants, transmission lines) that are described by a big set of diverse dynamically varying parameters. With no problem-oriented software, the above problems have been long solved by labor-consuming calculations done practically "manually" using the Microsoft Excel. Therefore, there emerged a necessity to create and use the software for complex scientific investigations aimed at designing and projecting the IPG expansion.

The accomplishment of the set goal required a technology (method) for non-structured data collection and processing. Such data involve isolated energy and power-related information collected from various sources, and algorithms for its conversion and storage. It was also necessary to integrate the collected information resources into a geo-information computing system with an object-oriented database.

We know only several close functional software analogs. These are the Asia Pacific Energy Portal by UN ESCAP [9], the Energy GIS-system developed by GEIDCO (China) [10], and APERC Energy Data Network Service [11]. All of them contain a considerable amount of data on power plants and transmission lines worldwide.

The above analogs, however, have no computing part, and in the context of this paper, may be termed only as geo-information systems. We have been developing both the Data Processing and Geo-Information Computing System in one software-product. Our software, besides visualizing the power data, is intended for energy experts (our users) to conduct optimization calculations and to construct well-developed projections for the IPG development, by using mathematical optimization model.

## II. THE DATA PROCESSING AND GEO-INFORMATION SYSTEM ISSUES

A research team of the Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences (ESI SB RAS) investigates the formation of interstate power grids. To boost the efficiency and quality of this research, we have developed the Data Processing and Geo-Information Computing System (DPGICS).

Figure 1 presents the main DPGICS structural-functional blocks and the data flows among them.

The functions of the Data Sources block are to collect data from heterogeneous information resources and to convert them to a uniform database structure used in the DPGICS. To store the information in a uniform structure

within the DPGICS, we have developed an Object-Oriented Database. The data represent a system or a set of files in a special machine-readable format containing the parameters for the database object, its text and numerical values with special separators.

The next DPGICS functional block is intended to operate the ORIRES optimization model. The DPGICS interface allows experts to adjust and vary the ORIRES input parameters, the number of nodes and constraints, to run the "optimizer," and to generate the final set of tables with the optimal solution results. Further, the obtained set of tables determining a certain scenario of the IPG development for the target year is also stored in the DPGICS data structure.

To visually represent and analyze the information from the object-oriented database in the DPGICS, we have developed the Data Representation Interface block. We use the maps built by the free Google Maps Application Program Interface (Google API) as a background on which semi-transparent layers representing various power parameters are imposed. This interface allows scaling the map and analyzing the trends of parameter variations over different time periods, for example, over the last 20 years. The user can see these trends on the built maps, and in plots or diagrams.

Further, we developed and run a special DPGICS block – Analytical Internet Service [12] to represent the investigation results on the Internet. At a request, the DPGICS-generated maps, tables, and diagrams (keeping some dynamic functions implemented through web programming) are exported into this web application. For example, the DPGICS web interface users can also analyze the power parameter variation dynamics on the maps for different years (only as simple examples with animation of these trends).

The DPGICS implementation involved the following technologies.

We used the Embarcadero Delphi XE, API Google Maps, SHP and XML files containing geographical and other energy & power data as a basis for construction of the DPGICS and the object-oriented database. To create the DPGICS web-interface, we applied the web technologies based on PHP, Ajax, JQuery, JavaScript, MySQL, and Bootstrap. The ORIRES optimization model is represented by a set of mathematical formulas and equations solved in the DPGICS through the general algebraic modeling system. The model parameters are set and written in/from the object-oriented database for their further processing and presenting in the DPGICS table-graphic interface.

The information technologies applied in the DPGICS development, allow full implementation of the stated problems. The under-construction DPGICS integrates in itself algorithms for collecting/loading and processing semi-structured data (in the stage of their loading and restructuring into an object form), and also technologies for graphical and cartographical data display. To present

Table 1. A fragment of data on the selected OODB object.

Parameter name	Parameter value
ID	170710110631
year_s	1980
lat	36.47750000
long	127.48083333
country	Korea_South
district	Chungcheongbuk-do
name	Daecheong_Powerplant
type_st	hydro
add_files	inet_170710110631.docx
n_rasp	90.00 90.00 90.00 90.00 90.00 90.00 90.00 90.00 90.00 90.00 90.00 90.00 ...
w_gen	240000.00 240000.00 240000.00 240000.00 240000.00 240000.00 240000.00 240000.00 240000.00 240000.00 240000.00 240000.00 ...
	...

the final information in the cartographical interface, we use the registered layers of GIS-maps with the basic components for the power infrastructure of the countries and interconnections.

### III. TECHNOLOGY FOR THE DPGICS DATA STORAGE AND PRESENTATION

As noted above, to store and process energy & power data, we have developed an object-oriented database (thereinafter – OODB) within DPGICS. The idea of the technology for data storage in the OODB is that all the information loaded/written in the system is structured through different program algorithms (it depends on a certain resource) and stored in the form of unique database objects describing power parameters of the real world entities. Forms of the OODB objects may be various and independent: power plants, transmission lines, region, country, power interconnection [12], and others.

Each OODB object represents a record containing an object unique identifier, a set of parameters describing it, and the values of each parameter in a text or numerical form, stored by year, month, day, etc.

The content of database objects is represented in CIS interface by dynamic editable tables. The table contains object parameters and its values, see Table I. The edited values are automatically recorded in an object form.

Further, a separate file is opened in a simple text editor, where user can see a fragment of parameters and values of the selected object, written in a special format. In the example above, there is a fragment of the data for the Daecheong Power Plant (South Korea EPS). The OODB stores about 100 power objects belonging to South Korea

EPS. Here we can see:

1. Text parameters of the object, i.e. the unique identifier (ID), year when the data collection started (year\_s), object name (name), latitude and longitude coordinates,



a region (country), a province (district), power plant/capacity type (type\_st), additional files describing the object (add\_file). There can also be photos, videos, etc. Some of them can be indicated as key field for aggregation, for example, one can aggregate objects by "type\_st", or by "country" etc.

2. Numerical parameters of the object, i.e. installed capacity of power plant for every year, starting in 1980 (n\_rasp, in MW), total electricity generation per year (w\_gen, in MWh) and any others.

This list of parameters is length unlimited and may occupy several pages. The inner structure of the OODB files is not only convenient for machine processing but also is clear to a person (human). The data structure universality and simplicity make them practically independent of the programmers.

The information storage technology in the object form has an advantage over relational databases for the following reasons.

- The number of records and their length in our data structure are not limited. Accordingly, one can store an infinite number of the object parameters and their output values in one string, for example, hourly load within an entire year or a period of years without creating additional relational tables.
- One may use a free type/format of parameter values, viz., fractional, integer, or text, stored in an unlimited range of values. For example, frequent change in the names of power entities by year can be presented in the form of the Name parameter, whose text values are written in one string with special separators. The values for such a parameter can be taken from the OODB for any requested year, or for some years, for example, to see the name change trend.
- There is no need for a large number of relational tables and indices when operating with big lists of different types of objects unconnected with one another.
- The necessity to compactly present comprehensive information on a power entity in one special-format file. Such a presentation provides confidence in the data integrity and excludes potential losses and data originating at complex SQL-queries "collecting" values for one or several entities from various relational tables.

We have developed program algorithms for data input/output from the OODB by using graphic and cartographic data representation. We developed program procedures to verify the data and check them for integrity. To exclude duplication of the objects collected from various sources, we have developed algorithms for comparison that employ simple methods of semantic and syntactic text analysis. It is also possible to visually search for the facilities (in the case of power plants) by geographical coordinates on the Google Maps: we revealed the cases, when the power plant geographical coordinates were intentionally or accidentally

"shifted" a few kilometers away from their real location. Most major objects in the OODB have real geographical coordinates and can be presented on the maps through the DPGICS cartographic interface.

Quality research requires maximum comprehensive information on the investigated entities or processes. In particular, to determine the current state of a power system in a region, it is necessary to have the information on the installed capacities of the given EPS. When searching and obtaining such information, a researcher faces subjective and objective difficulties, because information sources are, as a rule, separated and located on various Internet

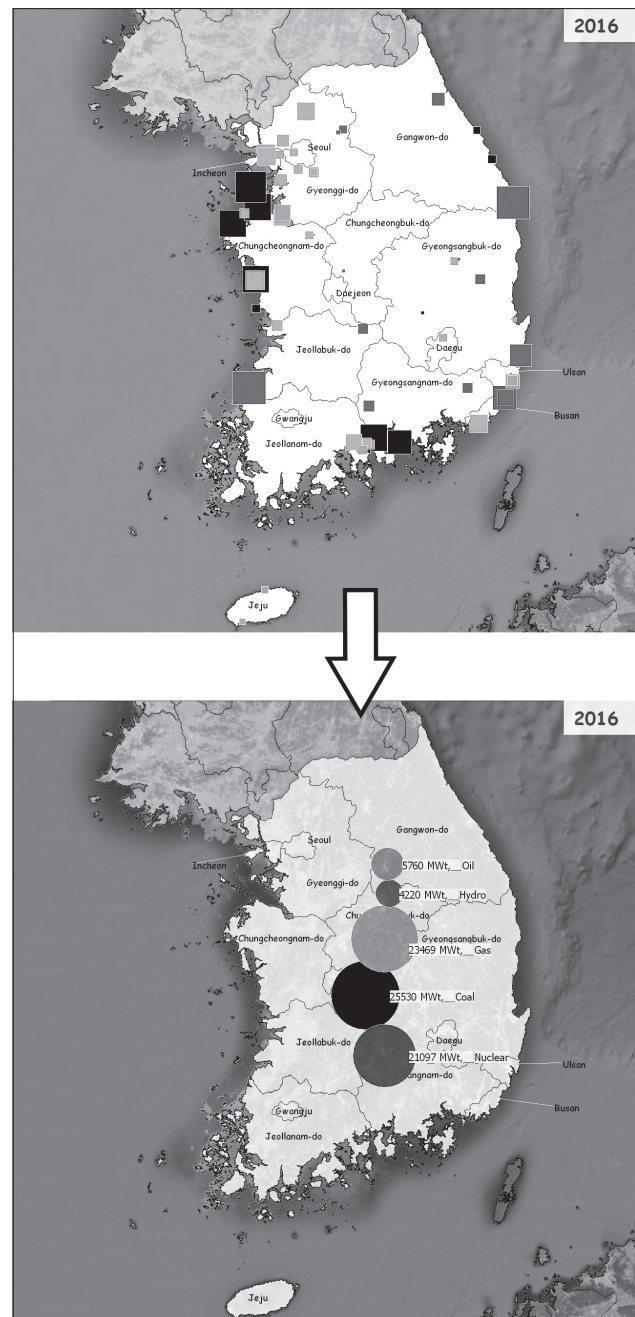


Fig. 2. Aggregation of power plant installed capacities in South Korea.

resources. To cope with this problem, we have developed a semi-automated method to collect/download semi-structured information from various sources and to convert it in an object structure with its subsequent verification and analysis. Each algorithm for downloading and converting the information in the OODB depends on certain types of data sources. We use combined (manual and semi-automatic) processing algorithms for data download such as parsing of MS Excel statistic tables, or automatic download of the entire data spreadsheets from various public energy databases like EnergyStorageExchange.org, EniPedia.org etc.

To load a great volume of the information presented in ontology form at DBPedia/EniPedia.org, for example, we have developed a program algorithm implementing the syntactic analysis of the XML-format data exported from the above Internet resource, and allocating the parameters of power plants and their values into the OODB structure [13]. To verify the objects downloaded from various sources in the DPGICS, we use simple algorithms to compare the obtained new objects with those already existing in the OODB.

This operation is executed for each new class of objects only once. In the event that there is updated information on the objects existing in the database (for example, new statistics for the current year), each program algorithm automatically distributes new values to the corresponding cells in the database structure. Based on the algorithms developed for various sources, we form and enlarge the OODB with unique information collected in a uniform structure whose usage expands manifold possibilities and enhances the quality of scientific research.

Further, the data from the OODB are processed through a convenient “query designer interface” (we have developed an add-in for SQL that uses the certain terms for convenient work of energy experts). This “query designer” allows us to extract object parameters, classify them, sample, and aggregate. The objects possess an expansive set of parameters that can be added when using the OODB. For power plants, in particular, we added the parameters for their cartographic binding and possibility of presentation on the DPGICS maps [14-15].

With the information on exact location of entities, the data can be aggregated at various levels, which opens up the possibilities for scaling the ORIRES model, i.e., for creating different levels of design diagram detailing. The collected data can be aggregated by country, region, national energy system or its subsystem. Figure 2 shows the aggregation (obtained in the DPGICS cartographic interface) of the power-plant installed capacities within the interconnected power systems (IPSS) of South Korea.

The levels of data aggregation can be different depending on the problem to be solved. The DPGICS automatically summarizes any object parameters for any year by the indicated key fields (by country ID, province, national power system, interstate power system, or by power plant type).

We formed a primary list of power plants in the OODB from the EniPedia.org open source. The power systems with their aggregated capacities (power plant types) shown in Figure 2, as well as individual power plants, are stored in the OODB like objects and their parameters.

The OODB object parameter values are also used as input data for the ORIRES optimization model being a

Table II optimal installed capacity (projected for 2035) by type of power plant (GW)

UPS	TPP			NPP	HPP	PSPP	WPP	SPP	TOTAL
	COAL	GAS	OIL						
Siberia (Russia)	25.09	3.58	0.00	2.40	28.62	0.00	0.00	0.00	59.69
East (Russia)	4.81	3.35	0.00	0.00	5.72	0.00	0.00	0.00	13.87
Sakhalin (Russia)	0.30	0.94	0.00	0.00	0.00	0.00	0.00	0.00	1.24
Mongolia	3.81	0.00	0.00		0.19	0.20	0.50	0.11	4.81
Northern China	489.00	20.30	0.00	24.10	3.20	25.00	35.00	24.00	620.60
North-East China	166.00	0.50	0.00	33.33	6.77	7.00	25.60	8.00	247.20
North Korea	3.00	7.00	2.00	0.00	7.01	0.00	0.57	0.71	20.29
South Korea	58.20	44.59	1.00	37.14	1.90	4.70	20.00	8.00	175.53
Japan	44.00	72.74	39.00	37.00	22.50	25.60	8.00	27.00	275.84

relevant functional part of DPGICS.

#### IV. VISUALIZATION OF CALCULATION RESULTS OBTAINED WITH THE MODEL

The ORIRES mathematical model for optimization of power system expansion and operating conditions (hereinafter - the Model) is used to calculate various scenarios for creation of interstate power grids and optimal expansion of interconnected power systems. All the Model indicated parameters are customized through a DPGICS interface. The DPGICS uses the general algebraic modeling system (GAMS) [16], which starts with the DPGICS interface, reads the Model parameters set in the interface, and forms output tables with an optimal solution.

The DPGICS graphic block enables the selected options to be visually assessed and interpreted on maps, plots, as well as spreadsheets to be built. High-dimensional matrices of primary output data are converted into aggregated table templates with the results obtained by the Model. The aggregated tables with prepared results of forecasting the expansion of power systems in the countries of Northeast Asia for 2035 are demonstrated as an example. Table II presents the structure of optimal installed capacity (projected for 2035) for various types of power plants. In this Table, TPP means thermal power plant; NPP – nuclear power plant; HPP – hydro power plant; PSPP – pumped storage power plant; WPP – wind-power plant; SPP – solar power plant.

Indicators from Table II (“total” column) are demonstrated on the map through the cartographic user interface, Figure 3.

The indicators can be displayed as pie charts on the map, for any region in the world. This map is created automatically through the DPGICS cartographic interface. In addition, the constructed maps can be exported to the DPGICS web application.

The DPGICS considerably simplifies the creation

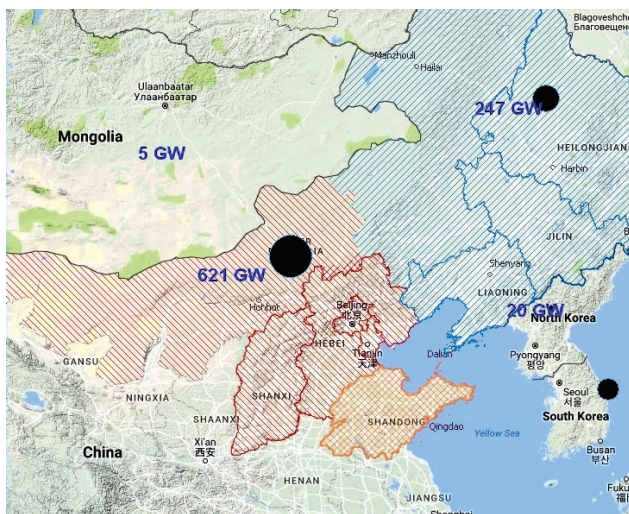


Fig. 3. Optimal installed capacities in EPSs of Northeast Asia countries.

of similar tables, reduces labor input when searching, checking, and forming the parameters for the optimization model, helps to visually represent the results of the obtained calculations in a graphic and cartographic form, and increase the research quality.

#### V. CONCLUSION

We have developed and upgraded a geo-information computing system to support scientific research into the formation of interstate power grids. Apart from visualizing the power data, the DPGICS allows experts to perform optimization calculations and to forecast the IPG expansion by using the ORIRES mathematical optimization model.

We propose different algorithms for data collection/download into the object-oriented structure, and an original method of data storing and processing in the object-oriented database, visualizing the calculation results obtained using the ORIRES optimization model in the aggregated tables, on plots and maps.

The projects for power cooperation between Russia and Eurasian countries will enable us to practically test the developed algorithms and information technologies to forecast the expansion of national and interstate power grids. Our information technologies are universal, and can be applied to investigate interstate power grids in various regions of the world.

#### ACKNOWLEDGEMENT

The study was supported by the Russian Foundation for Basic Research, grant No. 18-07-00495 A.

#### REFERENCES

- [1] A. A. Barsegyan, M. S. Kupriyanov, V. V. Stepanenko and I. I. Holod, “The Data Analysis Technologies. OLAP Data Mining”, BHV-Petersburg, 2007, p. 384 (in Russian).
- [2] V. A. Duke, “Data Mining”. Institute of Informatics and Automation Russian Academy of Sciences, Saint Petersburg, 2001, p. 150 (in Russian).
- [3] Omatsu Ryo. Interim Report by Asia International Grid Connection Study Group, 10th International Conference on Asian Energy Cooperation “AEC 2017,” *E3S Web of Conference*, Vol. 27. pp. 9-18, 2018.
- [4] M. Seung, “Role of Korea for Developing the East Asia Super Grid,” *20th International Conference on Electrical Engineering*. (Jeju, Republic of Korea, 15-19 June 2014), Jeju, 2014.
- [5] J. Yoon, D. Park, H.Y. Kim, “The Pre-feasibility Results of NEAREST between the ROK, and the DRPK, and RF,” *Proceedings of the 6th Intern. Conference “Asian Energy Cooperation: Forecast and Realities,”* Irkutsk, September 7-11, 2008, Irkutsk, pp. 59-67, 2008.



- [6] D. Chimeddorj, "Power Sector of Mongolia, Regional cooperation policy," *Northeast Asia Regional Power Interconnection Forum*, Beijing, China, October 26-27, 2016.
- [7] Tofael Ahmed, Saad Mekhilef, Rakibuzzaman Shah, Mithulananthan N., Mehdi Seyedmahmoudian and Ben Horan. "ASEAN power grid: A secure transmission infrastructure for clean and sustainable energy for South-East Asia," *Renewable and Sustainable Energy Reviews*, Vol. 67, pp. 1420-1435, 2017.
- [8] L. S. Belyaev, S. V. Podkoval'nikov, V. A. Savel'ev, L.Yu. Chudinova, "Efficiency of Interstate Electric Power Interconnections," Novosibirsk, Nauka, 2008, p. 240. (in Russian).
- [9] Asia Pacific Energy Portal: Interactive data and policy information. [Online] Available: URL: <http://asiapacificenergy.org/>, (access date 20.11.2017)
- [10] Xuming Liang, "Application and research of global grid database design based on geographic information," [Online] Available: *Global Energy Interconnection Development and Cooperation Organization*, Vol. 1, No. 1, pp. 87-95, 2018.
- [11] APERC Energy Data Network Service. [Online] Available: <https://www.egeda.ewg.apec.org/> (access date 11.03.2018)
- [12] ESAS: Energy Statistical Analytical Service. [Online] Available: <http://esas.com.ru/en> (access date 12.06.2018)
- [13] S. Podkovalnikov, I. Trofimov, L. Trofimov, "Data processing and optimization system to study prospective interstate power interconnections," In Proceedings of the 10th International Conference on Asian Energy Cooperation (AEC 2017), Irkutsk, *E3S Web of Conference*, vol. 27, pp. 32-40, 2018.
- [14] I.L. Trofimov, "Using metadata to query the database on thermal economy of Russia through the Internet, Mathematical and Informational Technologies," "MIT-2013," Vrnjacka Banya, Serbia, September 05-14, University of Pristina, pp. 706-714, 2013, (in Russian).
- [15] I.L. Trofimov, L.N. Trofimov, "Modern problems of search and verification of data on power plants in China and other countries, Informacionnye i matematicheskie tehnologii v nauke i upravlenii," Irkutsk, vol. 4, no. 8, pp. 120-128, 2017, (in Russian).
- [16] I.L. Trofimov, L.N. Trofimov, S. Podkovalnikov, "Data Representation from Energy Balances by Using Geo-information System", Vth International workshop "Critical Infrastructures: Contingency Management, Intelligent, Agent-based, Cloud Computing and Cyber Security," (IWCI 2018), Advances in Intelligent Systems Research. Atlantis Press. Vol. 158. pp.177-182, 2018.
- [17] M. R. Bussieck, A. Meeraus, "General Algebraic Modeling System (GAMS)," Chapter 8, J. Kallrath (eds), "Modeling Languages in Mathematical Optimization," *Applied Optimization*, Springer, Boston, MA, 2004, vol. 88.



**Ivan Trofimov** is an engineer and junior researcher at the Melentiev Energy Systems Institute Siberian Branch of the Russian Academy of Sciences.



**Sergei Podkovalnikov** is a Ph.D., head of the laboratory of Interstate Power Grids at the Melentiev Energy Systems Institute Siberian Branch of the Russian Academy of Sciences.



**Leonid Trofimov** is a lead programmer at the Melentiev Energy Systems Institute Siberian Branch of the Russian Academy of Sciences.



**Lyudmila Chudinova** is a Ph.D., senior researcher at the Melentiev Energy Systems Institute Siberian Branch of the Russian Academy of Sciences.