ATLANTIS PRESS
Advances in Intelligent Systems Research, volume 169
VIth International Workshop 'Critical Infrastructures: Contingency Management, Intelligent, Agent-Based, Cloud Computing and Cyber Security' (IWCI 2019)

Representation of Subject Knowledge from the Field of Vulnerability Analysis of Energy Systems in Distributed Applied Software Packages

Alexei Edelev
Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences
Irkutsk, Russia
flower@isem.irk.ru

Natalia Beresneva
Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences
Irkutsk, Russia
beresneva@isem.irk.ru

Sergei Gorsky
Matrosov Institute for System Dynamics and Control Theory of Siberian Branch of the Russian Academy of Sciences
Irkutsk, Russia
gorsky@icc.ru

Ivan Sidorov
Matrosov Institute for System Dynamics and Control Theory of Siberian Branch of the Russian Academy of Sciences Irkutsk, Russia
ivan.sidorov@icc.ru

Alexander Feoktistov
Matrosov Institute for System Dynamics and Control Theory of Siberian Branch of the Russian Academy of Sciences Irkutsk, Russia
agf@icc.ru

Abstract—The paper proposes a new high-level language of the Orlando Tools toolkit. It is designed for representing subject knowledge in studying the vulnerability of critical energy infrastructures using distributed applied software packages. The experimental results show that the applying of subject knowledge and information about the characteristics and use-rules of environment resources in Orlando Tools ensures us a number of advantages in comparison with other distributed computing management systems. Among them are selecting the optimal configuration of the computational infrastructure in solving large-scale practical problems, creating the effective problem-solving schemes, and the rational allocating environment resources, etc.

Keywords—critical infrastructure, energy system, vulnerability, distributed computing

I. Introduction

Energy systems are critical infrastructures because the disruption of their functioning adversely affects the economy of the state and the well-being of society. Functioning of energy systems during emergency situations in many respects depends on their vulnerability. The concept of vulnerability has two closely related meanings [1]. In the first interpretation vulnerability is seen as a global system property that expresses the extent of adverse disturbances caused by the occurrence of a specific hazardous event. In the second interpretation, vulnerability is used to describe a critical element of a system as the failure of that element causes large negative consequences to that system [2]. In this paper, the term vulnerability will be used to describe a system property in accordance with the first meaning.

Energy systems as critical infrastructures have such a property as interconnectivity [3-7]. Because of this property the energy systems can be integrated within the energy sector model.

Studying the operation of interconnected energy systems as critical infrastructures under emergency conditions requires the formulation of new problems that are based on a comprehensive analysis of all or selected combinations of various values of energy model parameters.

Usually, the total number of parameter value combinations for such problems is extremely large. This fact is a common feature of the vulnerability analysis of critical energy infrastructures. Thus, the combinatorial nature of such research justifiably requires the use of high-performance computing.

The subject knowledge description has a complex structure. To form such a description, languages of the ontological type can be used. For example, the Contingency Management Language (CML) is one of such languages [8]. Unfortunately, they cannot map models and algorithms of the studied subject domain to the architecture of a parallel or distributed computing environment.

In this regard, we propose a new high-level language for the subject knowledge representation in the vulnerability analysis of critical energy infrastructures field. This language is included in the Orlando Tools framework designed to develop and use distributed applied software packages [8] in a heterogeneous environment [10]. Nowadays, the language is used in developing packages that support solving a class of problems related to analyzing the vulnerability of energy systems. It ensures describing a fulfillment scheme (problem-solving scheme) of all stages of a computational experiment.

The paper is structured as follows. The next section briefly reviews the distributed computing management state in the energy security field. Section 3 represents the main subject domain aspects. Our approach to converting subject knowledge into the Orlando Tools computational model is considered in Section 4. Section 5 concludes the paper.

II. Related Work

Today, we can observe the high activity of studies related to the use of combinatorial methods for solving problems in the energy research field using high-performance computing. Zhang et al. [11] propose the parallel deterministic dynamic programming method and a hierarchical adaptive genetic algorithm to solve the problem, which involves many conflicting objectives and constraints in the context of reservoir system operation.

A fine-grained parallel discrete differential dynamic programming algorithm, which is based on Fork/Join parallel framework in multi-core environment, is proposed to improve the computing efficiency for long-term operation of multireservoir hydropower systems in [12].

Li et al. [13] develops a parallel dynamic programming algorithm to optimize the joint operation of a multi-reservoir system.

A novel efficient method called parallel progressive optimality algorithm for solving hydropower operation problem is presented in [14].

Feng et al. [15] proposes three parallel modes for multi-dimensional dynamic programming to solve the optimization problem of the cascade reservoirs operation. In addition, they provide the parallel multi-objective genetic algorithm for effectively solving multi-objective constrained optimization problem with two competing objectives and numerous physical constraints in water resource and power systems [16].

The description of the subject domain in each of the aforementioned studies is very specific. These descriptions are implemented within parallel programs that support single-level or two-level parallelism.

As shown in the paper [17], the computations speedup and the resource use efficiency decreases significantly when moving from parallel computing systems with shared memory to distributed computing environments. This is especially evident when using heterogeneous resources. Using traditional tools for managing distributed computing in the grid (for example, GridWay [18]) and cloud (for example, OpenStack [19]) does not provide an improvement to this situation.

Thus, the tools that enable us to associate features of problem-solving in the subject domain with the properties of the computing environment are relevant.

III. Subject Domain

The vulnerability analysis of critical energy infrastructures addresses the following problems:

- Assessment of the consequences of natural or technological extreme events for end-users of energy resources,
- Determination of the most vulnerable energy facilities (critical elements) in the event of these emergencies.

These interrelated and complementary problems are studied within the modeling scenario, supposed the formalization of energy infrastructure under study and its transformations in the methodology of the computational experiment implementation. The formalization of energy infrastructure results in a linear programming (LP) problem presented in the format understood by the external LP solver. Highly specialized software modules provide the implementation of the following operations:

- Setting up scenarios,
- Developing and optimizing the scenario cases,
- Transforming and visualization of the computation results.

Usually, the composition of these processing operations is different for various research tasks. The used (initial) and calculated (intermediate and result) data are represented by model parameters, calculation estimation indicators.

The above-listed software modules and categories of data and operations are described in the corresponding parts of the computational scheme defined by the subject domain ontology (in the blocks of parameters, operations, and modules). They are coordinated step by step within the problem formulation block (Fig. 1).

For the problem of critical elements detection of the energy sector, the computational scheme contains the energy sector state estimation operations according to the given model parameters, the failed elements criticality estimation operations, the element ranging operations based on the determined criticality estimations (Fig. 2).

IV. Converting Subject Knowledge into the Orlando Tools Computational Model

Orlando Tools provides the tools for a declarative specification of algorithmic knowledge, information needed for continuous integration of applied software of packages, data about software and hardware of environment nodes, and information about administrative policies in them. Algorithmic knowledge includes computational knowledge about modules for solving problems in the subject domains of packages, schematic knowledge about the modular structure of models and algorithms, production knowledge to support decision-making in selecting optimal algorithms for solving the problem. Orlando

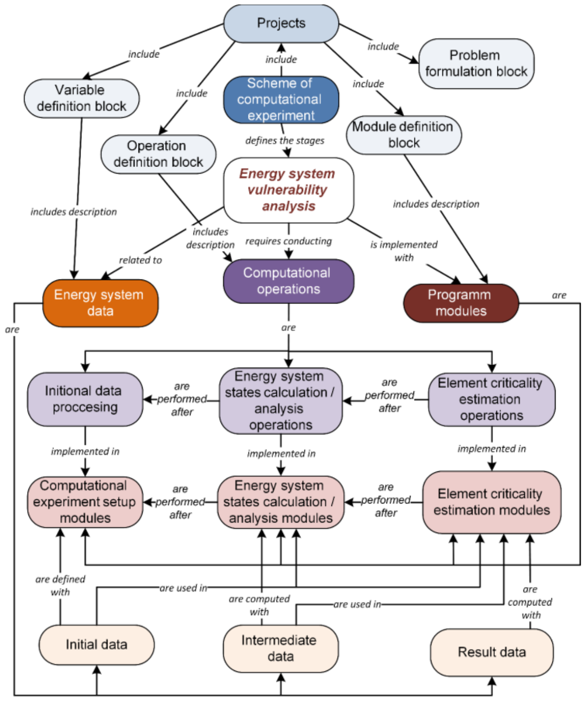

Fig. 1. The computational scheme defined by the subject domain ontology.

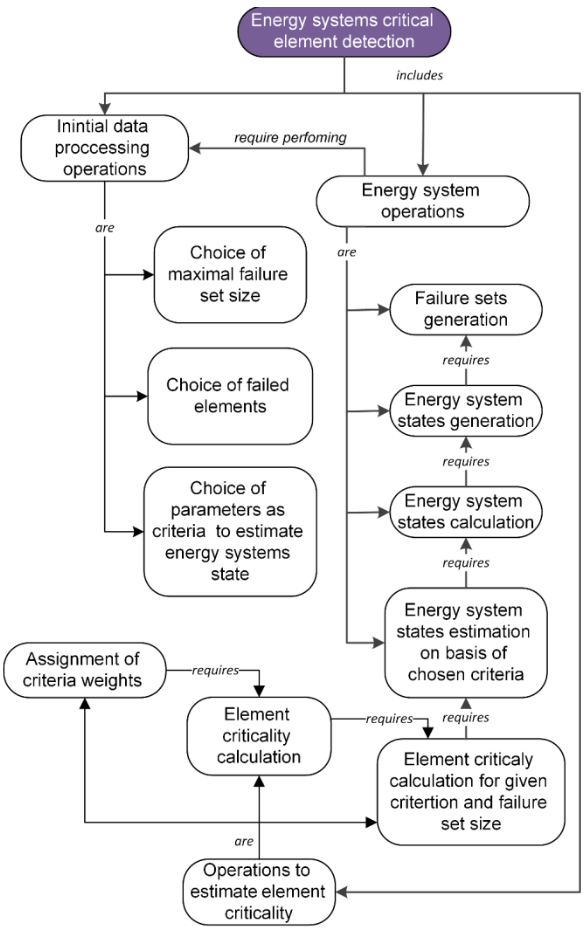

Fig. 2. The computational scheme for the energy system critical elements detection.

Tools uses XML notation as the main input language. It includes all the necessary information about a package. A large amount of this information and orientation to the XML description prevent an expert from comfortable work on this one.

In order to simplify the description of package computational model, a high level Computational Model Description Language (CMDL) has been developed for the subject domain experts. The CMDL syntax is shown in the Table 1. The key words **Project**, **Parameter**, **Operation**, **Module**, and **Task** correspond the following concepts of a subject domain: package, parameter, operation, module, and problem formulation.

TABLE I. THE CMDL SYNTAX

Syntax Item	Example
Project Project_name	**Project** R1
Parameter Parameter_name Extension;	**Parameter** z1 xml;
Operation Operation_name **module**|**task** Module_name|Task_name (Input_parameters> Output_parameters);	**Operation** o1 **module** m1 (z1, z3, z2, z4 > z7, z9);
Module Module_name (Input_parameters>Output_parameters) **Repository_name Project_name**;	**Module** m1 (data, elements, multiplicity, number_of_ subsets> subsets, statistic_file) repo1 R1.master;
Task t1 (Input_parameters> Output_parameters);	**Task** t1(z4, z1, z2, z3, z5, z6 > z8);

CMDL supports the parameters that are parallel lists of data, and operations over such parameters. The Orlando Tools interpreter for problem-solving scheme rationally allocates the heterogeneous resources to perform these operations taking into account both the existing and predicted computational load.

A tool to convert CMDL description to XML notation has been implemented with the PEG.js [20]. The converter allows not to change the Orlando Tools for developing packages but to enhance them with new features.

The computational model of the package for generation and calculation of the failure sets of a particular size is shown in Fig. 3. Information about the rules of calling modules is stored in a repository in specialized files. This allows the subject domain experts to describe a package easily. This information can be changed by means of the Orlando Tools editor. The example of the XML description generated by the converter is shown in Fig. 4. It is correspondent to the example in Fig. 3. We can see that the CMDL description with informative comments is more compact and clearer than the XML description without comments.

```
Project R1 //Generation and calculation of the failure sets of particular size
//Input parameters
Parameter l001 xml; //network scheme
Parameter l002 txt; // failure set size
Parameter l004 xml; //set of failed elements
Parameter l005 txt; //number of resources
Parameter l006 txt; //distributed database address
Parameter l008 xml; //list of filters as regexp expressions
//Temporary parameters
Parameter l100 xml [l005];//disturbance subsets
//Ouput parameters
Parameter l200 xml [l005]; // disturbance impact consequence subsets
Parameter l202 txt;//module m2 stat filename
Parameter l203 txt;// module m3 stat filename
Module m2 (scheme, failed_elements, size, number_of_disturbance subsets > disturbance_subsets [], stat_filename) repo1 m2.master;
Module m3 (scheme, failed_elements, disturbance_subsets, db_address, filters > consequence_subsets, stat_filename) repo1 m3.master;
Operation f2 module m2 (l001, l004, l002, l005 > l100, l202); // Generation of failure sets
Operation f3 module m3 (l001, l004, l100, l006, l008 > l200, l203); // Calculation of failure sets
Task t1(l005, l001, l002, l004, l006, l008 > l200); // Generation and calculation of the failure sets of size l002
```

Fig. 3. Description of the computational model of a package on CMDL.

The XML description of the computational model of the package is stored in the Orlando Tools knowledge base. Schematic knowledge of the package (the relationship between parameters and operations) can be represented in graphical form (Fig. 5).

In comparison with well-known distributed computing management systems, the experimental results obtained in the heterogeneous environment based on the resources of the Irkutsk Supercomputer Center of SB RAS [21] demonstrate a number advantages due to the use of subject knowledge and information about the characteristics and use rules of the environment in Orlando Tools [22-24]. We can highlight the following advantages:

- Selection of the optimal configuration for the computing infrastructure in solving large-scale practical problems,
- Creation of the effective problem-solving schemes,
- Rational allocation of environment resources,
- Improvement of the resources use efficiency,
- Increase of the computations speedup,
- Decrease of the problem-solving time.

V. CONCLUSIONS

The paper addresses the relevant problem related to associating features of problem-solving in the subject domain with the properties of the computing environment. Usually, subject knowledge is described using languages of the ontological type. Unfortunately, in this case, there is significant complexity with mapping models and algorithms of the studied subject domain to the architecture of a parallel or distributed computing environment. In this regard, we propose a new high-level language named CMDL. It is applied in the Orlando Tools framework that is intended for the development and use of distributed applied software packages in the heterogeneous environment.

CMDL supports the representation of subject knowledge in studying the vulnerability of energy systems and energy sector. Moreover, CMDL supports the describing the special parameters that are parallel lists of data, and operations above such parameters.


```
<pack>
 <project name="R1">
<parameter><name>l001</name><type></type><extention>xml</extention><list></list></parameter>
<parameter><name>l002</name><type></type><extention>txt</extention><list></list></parameter>
<parameter><name>l004</name><type></type><extention>xml</extention><list></list></parameter>
<parameter><name>l005</name><type></type><extention>txt</extention><list></list></parameter>
<parameter><name>l006</name><type></type><extention>txt</extention><list></list></parameter>
<parameter><name>l008</name><type></type><extention>xml</extention><list></list></parameter>
<parameter><name>l100</name><type></type><extention>xml</extention><list>l005</list></parameter>
<parameter><name>l200</name><type></type><extention>xml</extention><list>l005</list></parameter>
<parameter><name>l202</name><type></type><extention>txt</extention><list></list></parameter>
<parameter><name>l203</name><type></type><extention>txt</extention><list></list></parameter>
<operation><name>f2</name><input>l001,l004,l002,l005>l100,l202</input><condition>1</condition><while>0</while><type>module</type><module>m2</module><list></list></operation>
<operation><name>f3</name><input>l001,l004,l100,l006,l008>l200,l203</input><condition>1</condition><while>0</while><type>module</type><module>m3</module><list></list></operation>
<module><name>m2</name><parameter>scheme,failed_elements, size,number_of_disturbance_subsets>disturbance_subsets[],stat_filename</parameter><signatura></signatura><cores></cores><walltime>0</walltime><type></type><repo>repo1 m3.master</repo></module><module><name>m3</name><parameter>scheme,failed_elements,disturbance_subsets,db_address,filters>consequence_subsets,stat_filename</parameter><signatura></signatura><cores></cores><walltime>0</walltime><type></type>><repo>repo1 m3.master</repo></module>
<task><name>t1</name><parameter>l005,l001,l002,l004,l006,l008>l200</parameter><plan_type>0</plan_type><plan></plan></task>
 </project>
</pack>
```


Fig. 4. XML description of the computational model of a package.

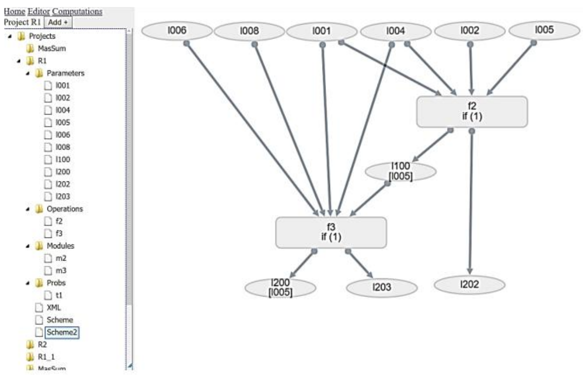

Fig. 5. Graphic representation of schematic knowledge.

Due to this fact, the Orlando Tools interpreter for problem-solving scheme can optimally allocate the heterogeneous resources to perform these operations. All aforementioned advantages of the proposed approach to the representation of subject knowledge were demonstrated in the process of carrying out large-scale computational experiments.

ACKNOWLEDGMENT

The research was carried out within the framework of scientific projects III.17.5.1 and IV.38.1.1 of the basic research program of the SB RAS. In addition, it was supported in part by the RFBR, project no. 19-07-00097-a.

REFERENCES

[1] J. Johansson and H. Hassel, "An approach for modelling interdependent infrastructures in the context of vulnerability analysis," Reliab. Eng. Syst. Safe., vol. 95, no. 12, pp. 1335–1344, July 2010.

[2] H. Jonsson, J. Johansson, and H. Johansson, "Identifying critical components in technical infrastructure networks," P. I. Mech. Eng. O.-J. Ris., vol. 222, no. 2, pp. 235–243, June 2008.

[3] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, "Identifying, understanding, and analyzing critical infrastructure interdependencies," IEEE Contr. Syst. Mag., vol. 21, no. 6, pp. 11–25, December 2001.

[4] E. E. Lee, J. E. Mitchell, and W.A. Wallace, "Restoration of services in interdependent infrastructure systems: a network flows approach," in IEEE Trans. Syst. Man. Cybern. Part C: Appl. Rev., vol. 37, no. 6, pp. 1303–17, November 2007.

[5] R. Zimmerman, "Social implications of infrastructure network interactions," J. Urban Technol., vol. 8, pp. 97–119, August 2001.

[6] D. D. Dudenhoeffer, M. R. Permann, and M. Manis, "CIMS: a framework for infrastructure interdependency modeling and analysis," in Proc. of the 38th Winter simulation conference, ACM, pp. 478–85, 2006.

[7] P. Zhang and S. Peeta, "A generalized modeling framework to analyze interdependencies among infrastructure systems," Transp. Res. Part B: Methodol., vol. 45, no. 3, pp. 553–79, March 2011.

[8] L. Massel and A. Massel, "Intelligent support tools for strategic decision-making on Smart Grid development," in E3S Web of Conferences, EDP Sciences, 2018, vol. 69, 02009.

[9] A. Feoktistov, R. Kostromin, I. Sidorov, and S. Gorsky, "Development of Distributed Subject-Oriented Applications for Cloud Computing through the Integration of Conceptual and Modular Programming," in Proc. of the 41st International Convention on information and communication technology, electronics and microelectronics (MIPRO-2018), IEEE, pp. 256-261, 2018.

[10] A. Feoktistov, S. Gorsky, I. Sidorov, R. Kostromin, A. Edelev, and L. Massel, "Orlando Tools: Energy Research Application Development through Convergence of Grid and Cloud Computing," Commun. Comput. Inf. Sci., vol. 965, pp. 289–300, 2019.

[11] Z. Zhang, S. Zhang, Y. Wang, Y. Jiang, and H. Wang, "Use of parallel deterministic dynamic programming and hierarchical adaptive genetic algorithm for reservoir operation optimization," Comput. Ind. Eng., vol. 65, pp. 310–321, February 2013.

[12] X. Li, J. Wei, T. Li, G. Wang, and W. W.-G. Yeh, "A parallel dynamic programming algorithm for multi-reservoir system optimization," Adv. Water Resour., vol. 67, pp. 1-15, January 2014.

[13] C. Cheng, S. Wang, K.-W. Chau, and X. Wu, "Parallel discrete differential dynamic programming for multireservoir operation," Environ. Modell. Softw., vol. 57, pp. 152-164, March 2014.

[14] Y. Zhang, Z. Jiang, C. Ji, and P. Sun, "Contrastive analysis of three parallel modes in multi-dimensional dynamic programming and its application in cascade reservoirs operation,"J. Hydrol., vol. 529, pp. 22-34, July 2015.

[15] Z. Feng, W. Niu, C. Cheng, and X. Wuvol, "Peak operation of hydropower system with parallel technique and progressive optimality algorithm," : Int. J. Elec. Power., vol. 94, pp. 267–275, July 2017.

[16] Z. Feng, W. Niu, and C. Cheng, "Optimization of hydropower reservoirs operation balancing generation benefit and ecological requirement with parallel multi-objective genetic algorithm," Energy, vol. 153, pp. 706-718, April 2018.

[17] T. Sharma, J. Glynna, E. Panosd, P. Deanea, M. Gargiuloe, F. Rogana, and B. O. Gallachoir, "High performance computing for energy system optimization models: Enhancing the energy policy tool kit." Energ. Policy, vol. 128, pp. 66–74, May 2019.

[18] GridWay Metascheduler. [Online]. Available: http://www.gridway.org.

[19] V. K. Bumgardner, OpenStack in action. Manning Publications, Shelter Island, 2016.

[20] Parser Generator for JavaScript. [Online]. Available: https://pegjs.org.

[21] Irkutsk Supercomputer Center of SB RAS. [Online]. Available: http://hpc.icc.ru.

[22] I. V. Bychkov, G. A. Oparin, A. N. Tchernykh, A. G. Feoktistov, S. A. Gorsky, and R. Rivera-Rodriguez, "Scalable Application for the Search of Global Minima of Multiextremal Functions," Optoelectron. Instrum. Data Process., vol. 54, no. 1, pp. 83–89, January 2018.

[23] A. Edelev, V. Zorkaltsev, S. Gorsky, V. B. Doan, and H. N. Nguyen, "The Combinatorial Modelling Approach to Study Sustainable Energy Development of Vietnam," in Communications in Computer and Information Science, vol. 793, pp. 207-218, November 2017.

[24] A. Feoktistov, I. Sidorov, A. Tchernykh, A. Edelev, V. Zorkalzev, S. Gorsky, R. Kostromin, I. Bychkov, and A. Avetisyan, "Multi-Agent Approach for Dynamic Elasticity of Virtual Machines Provisioning in Heterogeneous Distributed Computing Environment," in Proc. of the International Conference on High Performance Computing and Simulation, IEEE, pp. 909-916, 2018.