

OPINION

from Professor D.Sc. Ivan Garvanov
University of Library Studies and Information Technologies

Member of the Scientific Jury appointed by the Director of IICT-BAS via
№ 211/02.11.2020

SUBJECT: Dissertation of Kristina Ivanova Dineva with title “INTEGRATION OF HETEROGENEOUS DATA FROM DISTRIBUTED IoT DEVICES”, presented for the acquisition of educational and scientific degree “doctor” in a doctoral program “Informatics”, Professional field 4.6. “Informatics and computer science”.

1. General description

As a member of the Scientific Jury I have received:

- Dissertation
- Abstract in Bulgarian and abstract in English
- a list of printed scientific publications on the dissertation topic
- information about the fulfillment of the minimum requirements of IICT

2. Actuality, purpose and tasks of PhD dissertation

The integration of heterogeneous data from IoT devices is an extremely topical and scientific task, caused by the huge amount of heterogeneous data generated by different IoT systems and the need for this data to be stored, processed and analyzed in real time. The integration of disparate data necessitates different levels of competence in technological and scientific terms.

The aim of the dissertation is to develop a system and tools for processing, modelling and integration of heterogeneous data obtained by distributed IoT devices.

To achieve this goal, the following tasks are defined:

1. To propose a methodology for processing, modelling and integration of heterogeneous data obtained from distributed IoT devices.
2. To develop a modular hardware system architecture and to design an appropriate method for communication between distributed IoT devices.
3. To propose an architecture of a software platform and an approach for the organization of the services for the intelligent processing of heterogeneous data from an IoT system.
4. To construct valid machine learning models for experimental validation of the proposed methodology.

5. To provide possible application in intelligent agriculture of the proposed system and tools for integration of heterogeneous data obtained from IoT distributed devices.

3. Analytical characterization of PhD dissertation

The dissertation is structured in an introduction, five chapters, a conclusion, a list of publications on the dissertation, a list of observed citations on the dissertation, a summary of the results achieved and a bibliography.

Chapter one presents an analytical review is provided about the theoretical basis related to the problem area of the dissertation. It includes a short introduction, relevance of the topic, applications, challenges, and existing research solutions. The explanation is given for the need to build and apply a new methodology for working with heterogeneous data, which extends and improves the existing approaches.

Chapter two presents a systematic methodology for processing, modelling and integration of heterogeneous data obtained from IoT devices. A general conceptual schema of development is presented. At the next stage, the theoretical-methodological framework is described and explained in detail as a series of steps, grouped in four main stages. In the process of defining the stages, a review is performed together with a comparative analysis of the existing methods and approaches in identifying specific cases for which their application is correct.

Chapter three describes the process of architecting and developing an IoT platform. It consists of two dedicated systems for hardware and software, which are able to communicate with each other. The first section of the chapter presents the hardware architecture and an innovative method for IoT devices communication. The second section presents architectural solutions for software implementation of a server application based on microservices and implementation of a client web interface. During the development of the software system a new approach was used for the organization of the services responsible for the intelligent data processing and exchange in the IoT system.

Chapter four presents experimental implementation and validation of the developed methodology is accomplished. Two types of problems are identified, and solutions are provided for classification and regression analysis. For the purposes of the solutions, all the steps described in the methodology have been strictly followed. As a result of the performed experiment, validated models for machine learning are built, which are ready for integration in a production environment.

Chapter five presents a practical application of the developed system. The needs and benefits of its use are considered. A comparative analysis is done between the existing systems on the market. Based on this analysis, a comparative characteristic has been compiled, summarizing the usefulness of the existing systems in relation to the developed IoT system in this dissertation.

The Conclusion presents a summary of the results obtained from the development. Guidelines for future research and improvement have been identified. A list of scientific publications on the dissertation and noted citations are presented.

The dissertation consists of 166 pages, 55 figures, 16 tables and 175 bibliography sources.

4. Contributions

I accept all the doctoral student's contributions, namely:

1. A methodology for processing, modelling and integration of heterogeneous data obtained from distributed IoT devices has been developed along with performed selection for:

- methods for working with heterogeneous data.
- machine learning classification and regression algorithms.
- metrics for evaluation and validation of obtained results.

2. An architecture of a modular hardware system has been developed which consists of sensors and IoT modules for control and communication. An innovative method of communication between IoT devices on the base of hierarchical IP addressing has been proposed.

3. An MSA software architecture for storage, processing and analysis of heterogeneous data has been designed and implemented. An innovative approach for the organization of the services for the intelligent data processing and exchange in the IoT system has been developed, which increases the reliability and functionality of the system and additionally provides capabilities for machine learning.

4. Machine learning models have been trained and build, which experimentally confirm the developed methodology.

5. A possible application of the developed IoT system for integration of heterogeneous data in intelligent agriculture is shown. A comparative analysis of the functional characteristics and market positioning of existing similar systems is done, through which the economic efficiency and expediency of the developed IoT system are proved.

5. Abstract and author' declaration

The presented two versions of the abstracts in Bulgarian and English correctly reflect the content of the dissertation and correspond to the requirements of Bulgarian legislation. From the presented declaration of originality, as well as from the presented papers on the dissertation theme, it can be judged that the described results are a personal work of the author.

6. Assessment of compliance with the minimum national requirements

The PhD student *Kristina Dineva* has tested parts of his dissertation in eight scientific publications, all in English. One of the publications is with Q2, one is with Q3, four are with Q4 and two are indexed in Scopus.

According to the minimum national requirements for obtaining of the educational and scientific degree „Doctor in the professional field 4.6 " Informatics and computer science ", the required scores are to be at least 30 for the group of indicators G. The same number of scores is required by the Regulations on the Conditions and Procedures for Acquisition of Academic Degrees and Occupation of Academic Positions in BAS and the Regulations on Specific Conditions for Acquisition of Academic Degrees and Occupation of Academic Positions in IICT-BAS. The presented publications on the dissertation form a total 190 scores for the indicators from G group, which is significantly higher than the required minimum of 30 scores. Eight of the publications are cited in 19 scientific publications.

7. Critical remarks and recommendations

The formulas in the dissertation are not numbered and it is difficult to cite them in the text.

Some of the described results in the dissertation are not reflected in publications and it is recommended to be published.

8. Conclusion

The presented dissertation meets the requirements of the Law on the Development of the Academic Staff in the Republic of Bulgaria, the Rules for its implementation, the Rules for the conditions for the acquisition of academic degrees and for the occupation of academic positions in the Bulgarian Academy of Sciences and the Rules for the specific conditions for the acquisition of scientific degrees and for academic positions at IICT-BAS.

The described results in the dissertation, along with the fulfillment of the national minimum requirements, give me enough reason to give a positive assessment of the dissertation work and **I suggest to the honorable scientific jury to award to Kristina Ivanova Dineva the educational and scientific degree „Doctor” in doctoral program "Informatics”, professional field 4.6 "Informatics and computer science".**

20.11.2020