REVIEW

about the PhD thesis for acquisition of the scientific degree **“doctor”**

in the professional field **4.6. „Informatics and computer sciences“**

Ph.D. program **„Informatics “**

Author of the PhD thesis**: Rosen Petrov Hristev**

Title of the PhD thesis**: Recovery of information arrays in a cloud environment**

Ph.D. supervisor: Professor Angel Golev, PhD

Reviewer: Prof. D.Sc. Ph.D. Eng. Todor Atanasov Stoilov, Institute of information and communication technologies – Bulgarian Academy of Sciences (IICT-BAS), Sofia, Acad.G.Bontchev str., BL.2

1. General notes

The Ph.D. thesis contains 117 pages. It is structured in 3 chapters, contributions, list of published publications, list of references. The list of references is not numbered and the reviewer did not find it necessary to count the presented bibliography.

2. Actuality of the problems in the PhD thesis

The contribution of the Ph.D. thesis refers to the development of software solutions for the recovery of deleted and/or overwritten data as well as previous versions of deleted files. This allows files that have been maliciously encrypted or deleted to be recovered as well as their previous versions. This thematic orientation of the Ph.D. thesis is important for the implementation of security over digital data that is processed in computer systems. Due to the presence of malicious software that "corrupts" computer system data, the Ph.D. research seeks and develops solutions to restore deleted files, encrypted files and the previous versions of working files. The scientific-applied result of the Ph.D. work is expressed in the development of program (script) solutions for restoring such "damaged" files and restoring them to their initial state. The Ph.D. research applies these solutions to protect data files in system what the student calls "cloud" computing environments.

I assess positively the thematic orientation of the Ph.D. works. I find that the research described in the Ph.D. thesis has a scientific-applied, pragmatic and useful value, as it is addressed to necessary and important elements of problems of effective management and access to general data in computer systems. I think that the topicality and importance of the developed problems are easily understandable.

The reviewer assesses the researches in the Ph.D. thesis positively. The scientific and applied part of the dissertation research is easy to understand. I think that the importance of Ph.D. research is evident, the results obtained are useful and this gives a positive certificate for the qualification of the candidate.

3. Degree of knowledge of the state of the problem and academic interpretation of the literary material

The Ph.D. work makes a meaningful presentation about the problems that have a place in the storage and general use of data in computer systems. A range of threats to information arrays that can lead to data loss are reported.

In chapter 1, an analysis of data storage methods was made. Accordingly, the most common threats that lead to data loss are presented. Data storage is considered in a "cloud" computing system architecture. In terms of content, user data is stored on remote computers to which users have remote access. This chapter analyzes publicly available "clouds" where users can store data: Google Drive, DropBox, OneDrive, Pcloud, Amazon Drive, JustCloud. The essential part of the analysis in this chapter is oriented towards the ways of accessing the information arrays, which is a prerequisite for implementation and support of the security of the information arrays and data. Accordingly, the types of threats that an information structure may encounter during its operation are evaluated. Malware activation, content and results are also commented. The inevitability of developing new data storage methods that have common use by multiple users has been noted as the intention of the research.

4. Correspondence of the chosen research methodology and the set goal and tasks of the dissertation with the contributions achieved

The aim of the Ph.D. thesis was to develop methods for the collective use of information resources in a distributed "cloud" computing environment. These methods should be implemented and applied for the efficient usage of general data and for protection against malicious software. In Chapter 2, an analysis of ways of joint use of file arrays by multiple users is made. The usage of these arrays involves editing, modifying, and deleting operations. The Ph.D. thesis purposefully examines the need to manage access and use of shared files, but in a technological environment that functions not as a local network, but as a global one. In a global environment, the work of program clients with common file resources is performed with a web browser, through a synchronization client and by applying the WebDAV protocol. The dissertation gives preference to the latter solution. In-depth knowledge of the use and configuration of this protocol in various versions of the Windows operating system is demonstrated, and potential problems in its use are commented on. As a goal of the second chapter of the dissertation work, I see the intention of the PhD student to implement a "cloud" organization for multiple uses of file systems implemented in an existing "private" computer infrastructure system. This makes it possible to change working with files on a local network as a way of working in a "cloud" (distributed, remote) computer system. For this purpose, attempts have been made to quantitatively estimate the required volumes of information space from the point of view of recovery of deleted or maliciously encrypted files. The good knowledge about the working processes of sharing files between multiple clients has allowed the PhD student to define a sequence of actions for moving to a "cloud" sharing procedures.

This sequence is named "author's cloud deployment method". It is noted that it is implemented in the practice of several organizations.

The reviewer positively assesses the applicant's qualifications in the area of analysis and modification of file system access methods. He would like to see that it has been given data about the evaluation of the effect of such implementation of a "cloud" solution or by comparisons between indicators of traditional and cloud based file system about abilities to restore lost data. Currently, the presentation in chapter 2 of "author's method" has a declarative form, but has not been evaluated in terms of usefulness, functionality and/or other criteria.

Chapter 3 develops and presents practical solutions for recovering deleted or overwritten data. Overwritten data is understood to restore previous values of already changed data. The PhD student again demonstrated deep knowledge about the processes if erasing and encryption by malicious actions of a software virus takes place. The rationale for data recovery stems from the ability to use the file directory of a user's deleted files. This directory is the data recovery source. Chapter 3 again defines a sequence of actions that determine how to recover data. This sequence is called the "author's method" RDSC. This chapter presents scripts that were implemented to recover data on a Widows station infected with CERBER crypto virus; recovering data from a LINUX operating system infected with the GonnaCry virus; recovery of database records and application software saved on Windows Server infected with Sodinokibi crypto virus; and same types of data from databases in a Linux operating system environment again infected with GonnaCry. This chapter presents and develops scripts for recovering deleted data stored in a cloud environment; scripts to restore previous versions of files.

The results of chapter 3 once again demonstrate the high professional experience of the doctoral student in the field of analysis of the ways of operating information flows in software operational environments. This has allowed, due to the knowledge of the internal processes of information interaction, to find solutions in case of data loss or malicious encryption and to successfully restore them. The reviewer expresses his opinion that he would like to see an evaluation of how effective these developed scripts are or to be compared with similar data recovery solutions.

5. Scientific and practical achievements in the PhD thesis

In the Ph.D. thesis modifications are made to a sequence of actions that allow to "clean" a computer system from a malicious virus and, as a result, to restore again the arrays of data of file systems, database, user programs. These sequences are referred in the thesis as "author's methods". The reviewer appreciates the pragmatic value of these "methods". An essential element in them is the developed scripts that implement functions for restoring the source data. The scripts are targeted for predefined case of loss and for respective operating system. The reviewer assesses that the scientific-applied contribution of the dissertation work is found in these scripts developed by the author. This contribution is an upgrade over technologies that apply to multiple access to data files and which I appreciate as a positive outcome of the Ph.D. work.

I find that the developed topic has a scientific-applied nature in the part of developing the author's scripts for data recovery after malicious intervention. I positively assess the results of the doctoral student's research. They have a useful practical character and potential for practical application by proving the usefulness and pragmatics of research in the dissertation work.

I consider these contributions to be sufficient for this Ph.D. work. They prove that the doctoral student can independently carry out research activities, solve problems that have a complex and difficult nature due to the complex and implicit nature of processes in information environments. These studies have an important place and a positive potential for the current management of virtual information systems.

When reading the dissertation work, I am convinced that the achieved results are mainly done personally by the Ph.D. student.

6. Correspondence with the minimal national legislative requirements

The reviewer assesses that the submitted publications correspond to the topic and content of the dissertation work. Three publications are presented. Two publications have been presented at conferences in India and Turkey and have a digital DOI registration. One publication is in a journal in our country which has Scopus rank SRJ=0.14, quintile Q4.

Participations in conferences in our country, AMEE-2021, Sozopol and IMEA-22, Plovdiv are also declared, but these participations in the documents of the doctoral student are not presented with relevant publications.

My assessment is that the presented publications are representative and satisfy the requirements for the defense of the educational and scientific degree "doctor".

The PhD student does not provide a list of citations of his publications. I appreciate that the presented publications are substantial and give a positive certificate for the doctoral student for his obtained results.

According to the legislative rules for defending the Ph.D. level for covering the minimal national requirements in professional field 4.6. "Informatics and computer sciences", it is required at least 30 points in Group G. The submitted documents for this dissertation present a reference with the needed data. My personal checks of the submitted publications prove that the presented publications satisfy this requirement by criteria G7: scientific publications referenced and indexed in world-recognized databases (Web of Science and SCOPUS).

7. Significance of the research and application achievements in the PhD thesis

The Ph.D. student Rosen Hristev demonstrates skills for analysis and evaluation of information processes that have a place in complex computer systems. These skills and knowledge are aimed at finding solutions to overcome the harmful results of malicious impact on information processes. The practical utility of the research lies in the recovery of deleted data, malicious encryptions. The good qualification of the doctoral student is also proven by his proposed solutions for data recovery for different operating systems, when affected by various malicious crypto viruses. Practical use of scripts developed by the PhD student to overcome the actions of crypto viruses has been declared.

The PhD student demonstrates skills in the analysis of information processes that take place in complex and difficult to identify settings.

The reviewer assesses that the Ph.D. research is useful and has led to potentially pragmatic outcomes such as finding solutions for data recovery, previous versions of data, overcoming crypto virus attacks.

In the presented documents, there is no data on the distribution of author contributions in the presented publications.

8. Few assessments, recommendations and remarks

I positively assess the presented Ph.D. work. It is evident from its content that the doctoral student has conducted independent research work.

The reviewer has no negative comments regarding the content of the thesis.

As a form of remark is the question of evaluating properties, qualities of developed scripts. Currently, the thesis applies declarations that the proposed solutions are useful, they have worked and they have been applied in various organizations. The reviewer assumes that an academic, scientific study should illustrate a benefit, an effect that can be evaluated by comparison with analogous solutions. Therefore, I recommend the doctoral student in his future work to use comparative evaluation methods to make sure other potential users about his solutions that they bring a greater positive effect.

Due to difficulties of terminological and abstract nature of the presented thematic area, I recommend the doctoral student to apply more block diagrams when explaining the processes in information systems. The graphical way of presentation brings a greater understanding of the content of the description compared to a pure textual presentation.

The reviewer considers that the doctoral student Rosen Hristev shows experience and qualifications for conducting independent research in the field of analysis and modification of the modes of operation of information processes and systems.

Conclusion

I give positive assessment about the scientific-applied and applied results in the PhD thesis of Rosen Petrov Hristev. I found that the legislative requirements of the Law for academic growth in Bulgaria, the Regulations for its application are satisfied.This gives me the reason to recommend to the honorable Scientific Jury to award **Rosen Petrov Hristev** the educational and scientific degree "Doctor" in professional field 4.6 "Informatics nd computer sciences”, scientific specialty "Informatics”

6.12.2023 Reviewer:

Prof. D.Sc. Ph.D. Eng. Todor Stoilov