

## Associate professor PhD István-Gergely Czibula

Summary of Habilitation thesis

### Contributions to Search-based Software Engineering

The present habilitation thesis presents a part of the candidate's research activity after defending the PhD Thesis at the Babeş-Bolyai University of Cluj-Napoca in December 2009. The activity presented in this thesis represents the candidate's original work and is highly relevant for his scientific and academic career, being an argument for the future independent development of his research and teaching career.

The main research direction of the candidate since 2010 is in the field of *Search-based software engineering*, continuing the research started in 2006 within the PhD studies. *Search-based software engineering* is an interdisciplinary research and applicative domain in which the focus is to formulate the software engineering problems as search ones and to apply computational intelligence techniques for solving them.

After the completion of the PhD thesis, the candidate continued to approach important software engineering problems, the main goal being the development of novel methods for automating different activities related to the maintenance and evolution of software using computational intelligence techniques. The problems approached are of major importance during software development and are complex, which is why computational intelligence techniques were developed to find good solutions to them. Since maintenance and software evolution problems are of utmost importance to the modern programmer, there exists a grown interest in the automation of as many processes in the lifecycle of software as possible, as well as the development of adequate mathematical models.

Since 2010, several research directions towards using machine learning techniques in the resolution of practical problems in software engineering have been proposed, such as: *software restructuring*, *behavioral adaptation* of software systems, *software defect prediction* and *adaptive relational association rule mining*. The proposed approaches are interdisciplinary and have contributed to the fundamental and applicative research in the domains of Applied Computational Intelligence and Software Engineering. They offer solutions based on machine learning for the most important problems of Software Engineering. The problems tackled have a major practical importance since software developers face them daily, therefore the development of automated techniques which offer solutions to these problems would lead to more correct and error-free software. Since it is very hard to identify direct solutions, classification methods based on machine learning are extremely useful in solving the aforementioned problems. The theoretical research results obtained were used to develop software frameworks which included all the computational techniques developed and are useful to assist the software developers in the maintenance and evolution stages of the software lifecycle.

The previously mentioned scientific results were disseminated through **80** scientific papers and 1 book in the search-based software engineering field. Among these publications 39 are indexed in ISI Web of Knowledge (19 are in Science Citation Index Expanded and 20 are in Conference Proceedings Citation Index - Cumulated Impact Factor=**30.201**) and 35 papers are indexed in prestigious international databases (SCOPUS, INSPEC, IEEE, ACM, DBLP, Mathematical Reviews). Among prestigious ISI journals in which the candidate has

published the research results, are mentioned: *Information Sciences, Knowledge and Information Systems, Applied Soft Computing*.

The research results obtained so far by the candidate had a large impact on the fundamental research (novel research directions in the search-based software engineering field were opened: unsupervised machine learning for refactorings identification, unsupervised machine learning for aspect mining, supervised machine learning for behavioral adaptation of software systems). The international visibility of the scientific activity is sustained by more than **183** citations. A large academic impact was among the young researchers and PhD students from our country and from abroad, proven by citations in recent PhD thesis, Master Thesis and Technical reports. All of these citations in prestigious journals (66 ISI citations) and theses confirm the international visibility of the candidate's scientific activity. Among the prestigious ISI journals in which the research results were cited, are mentioned: *Information Sciences, Knowledge and Information Systems, Applied Soft Computing, Applied Intelligence, Information and Software Technology, IEEE Transactions in Software Engineering*.

Since 2006, the candidate participated in six research projects. He was the principal investigator in two national research projects, scientific executive in one national research project, member in other two national research projects and member in one international project. All these research projects are in the field of *software engineering* and *computational intelligence*.

Within the research group on applied computational intelligence, the candidate has worked with three former PhD students and contributed to the finalization of their PhD thesis. Also, he are currently working with three PhD students and have already collectively published in prestigious ISI journals.

The thesis is structured as follows.

Chapter 1, **Introduction** presents the candidate's main scientific, professional and academic achievements, emphasizing the international visibility of his research activity.

Chapter 2, **Scientific achievements** presents the main scientific results which were obtained starting from 2010. The relevant original contributions are in the field of *search-based software engineering*, particularly in four main directions: *software restructuring, behavioral adaptation of software systems, software defect detection* and *relational association rule mining*. In the direction of *software restructuring* unsupervised learning based methods were introduced and the problem of adaptive restructuring was also investigated. Two supervised learning models for dynamically selecting the most suitable data structure as implementation for a data container were developed in the direction of *behavioral adaptation of software systems*. The problem of *software defect detection* was investigated using relational association rule mining, as well as using *crisp* and *fuzzy self-organizing maps*. The relevant original contributions in developing novel machine learning based techniques are also presented in Chapter 2. In this direction, a reinforcement learning based software framework for solving combinatorial optimization problems, a software framework for analyzing object oriented software systems, as well as the novel concepts of *adaptive* and *fuzzy* relational association rules are presented.

The future development plans from the scientific, professional and academic viewpoints are presented in Chapter 3, **Scientific, professional and academic future development plans**.

The thesis ends with the list of bibliographical references used over the thesis.