



FACULTY  
OF MATHEMATICS  
AND PHYSICS  
Charles University

## Report on the PhD thesis by Přemysl Čech

Nowadays, large collections of images/videos are created and stored for personal, medical, scientific, and industrial purposes. In order to benefit from this rich data source, models enabling effective and efficient content analysis, management and search are required. These models can be divided to machine learning based approaches relying on large training datasets and similarity search approaches depending on domain expert knowledge.

The proposed thesis investigates models for content-based similarity search, mostly using the metric space approach. More specifically, the approach assumes that the data are transformed to a descriptor universe (e.g., a vector space) with a metric distance that models similarity of descriptors. The candidate presents several content-based approaches for searching and interactive exploration of multimedia collections, including prototype demo applications. In addition, similar search and exploration principles are tested also for the network security domain.

In the last three years, the candidate focused also on new methods for efficient evaluation of distributed similarity joins that can be employed in presented content-based search and exploration applications. Using experimental evaluations, the candidate demonstrated that the exact variant of distributed similarity joins can be effectively and efficiently approximated with a heuristic relying on a set of reference points and Voronoi partitioning. These heuristics were also compared with a method providing guarantees.

The thesis is written in a readable introductory form, presenting basic terminology, showing example applications from various domains, and then presenting the main contributions. The presented contributions are based on candidate's publications, where the main chapter on similarity joins is taken from the recently accepted journal publication. The list of candidate's publications comprise 16 publications in peer-reviewed international conferences (including 1x CORE A\* and 3x CORE B conferences) and three impacted journal publications. The number of citations of his publications reaches 43 on scopus (without self citations) and his h-index is 4.

Based on the results and evaluation above, I **recommend** Přemysl Čech to obtain the Ph.D. degree.

10.1.2020

RNDr. Jakub Lokoč, Ph.D.