Tartu University
Faculty of Science and Technology
PhD projects 2017
Topics & Abstracts

January 14, 2017

Contents

1 Institute of Computer Science 2
  1.1 Massive open online courses in programming: the Case of Estonia . . . 2
  1.2 Business Process Privacy Analysis and Optimization .................. 2
  1.3 Multi-Perspective Declarative Process Mining ....................... 2
  1.4 Parallel privacy preserving computations ............................ 3
  1.5 Practical protocols for quantum position verification ............... 3
  1.6 Deep neural networks for microscopy images ....................... 3
  1.7 Secure E-Voting ............................................. 4
  1.8 Ensemble learning under context changes ........................... 4
1 Institute of Computer Science

1.1 Massive open online courses in programming: the Case of Estonia

Supervisor(s): Reelika Siwiste

The aim of this doctoral project is to investigate the development and impact of the programming MOOCs in Estonia and the associations between participants motivation, accomplishment and usage of support mechanisms in the courses. The main outcomes will be: (I) a research-based overview of the development of the programming MOOCs in Estonia; (II) developing the tool for analysing motivation in programming MOOCs; (III) the extent to which the accomplishment in the programming MOOCs is affected by the support mechanisms and differs in several courses using helpdesk; (IV) give an overview of the impact of the programming MOOCs in Estonia. After the study, it will be possible to give suggestions for educators and policy-makers how the use of programming MOOCs can be a considerable part in future curricular innovation.

1.2 Business Process Privacy Analysis and Optimization

Supervisor(s): Raimundas Matulevičius

The aim of this doctoral project is to develop and evaluate techniques for privacy analysis of business processes. The main outcome will be a tool that takes as input process models with privacy metadata, and analyses these process models in order to: (i) detect unintentional disclosures of private information; and (ii) quantify the amount of private information leaked by the outputs of the business process. The tool will generate reports that explain to data owners the maximum extent of possible leakage of private data. The tool will also suggest possible counter-measures to reduce privacy leakages in a business process as well as the specific points in the business process where these counter-measures should be deployed.

1.3 Multi-Perspective Declarative Process Mining

Supervisor(s): Fabrizio Maria Maggi

The aim of this doctoral project is to develop and evaluate declarative process mining techniques for analysing business processes from multiple perspectives (e.g., data, time, resources). The first step will be the definition of a graphical language based on a formal semantics that allows for the definition of multi-perspective declarative rules. Starting from this language a tool will be implemented (possibly available on the cloud) that takes as input an event log and/or a set of multi-perspective business rules, and analyses these inputs in order to: (i) discover a set of rules from an event log to describe the behaviour of a business process as recoded in the log; and (ii) monitor at run-time streams of events
against a set of existing multi-perspective declarative rules. The tool will generate reports that explain to data owners the behaviour of the underlying business process. The tool will also suggest possible actions to improve the performances of the business process as well as the specific points in the business process where these actions should be deployed.

1.4 Parallel privacy preserving computations

Supervisor(s): Eero Vainikko, Peeter Laud, Benson Muite

Novel privacy-preserving protocols for parallel graph algorithms will be examined and implemented to allow for fast privacy preserving analysis of data. Parallelization techniques that bring together linear algebra algorithms and classical graph algorithms to give faster execution time while still maintaining privacy will be examined. In addition communication hiding and communication avoiding protocols will also be examined, to allow for privacy preserving techniques on data in computer systems at separate sites.

1.5 Practical protocols for quantum position verification

Supervisor(s): Dominique Unruh

Quantum position verification (QPV) refers to a class of protocols in which the position of a mobile device is determined in space through the use of both quantum mechanical and relativistic effects. A QPV protocol ensures that the determined location is correct even if the mobile device is malicious. Thus QPV protocols allow us to control the access to a resource using spacial location. (E.g., access to a resource is only permitted within a certain building.) It has been shown by Buhrmann et al. that controlling access in this way cannot be achieved without using quantum mechanics. A number of QPV protocols have been proposed, most recently by the Unruh. However, the protocol by Unruh, as well as all protocols before, rely on strong technological assumptions (such as perfect quantum communication). Furthermore, the protocol is also based on strong cryptographic assumptions (such as the existence of random oracles).

The goal of this project is to bring QPV protocols closer to practical realizability by designing better protocols and developing new analysis techniques. The main focus is on QPV protocols that use cryptographic assumptions (i.e., we assume an attacker that is computationally limited). This is because Buhrmann et al. showed that the task is impossible without any limitations on the adversary.

1.6 Deep neural networks for microscopy images

Supervisor(s): Leopold Parts, Raul Vicente Zafra

Deep neural networks are an exciting class of models for high-dimensional large-scale data. Their promise is to leverage the very large scale of training information to learn networks with millions of parameters that capture the relevant structure for a particular
task. This is especially important in complex application domains, such as image analysis of high throughput microscopy, where the data are abundant, but there is no simple mathematical model to describe the signal for common statistical approaches.

Here, we propose to develop deep neural networks for high throughput microscopy data. A good model will deliver a hierarchical feature representation of the cells in the image, from dots, edges, and puncta, to larger organelles. The models will be trained for segmenting cells and subcellular compartments on a per-pixel basis, as well as end-to-end to describe entire images. Clonal cell population images from basic biological assays will be modeled to accelerate scientific discovery in collaboration with academic and industrial partners, and same principles will be applied to histology images from clinical collaborators to improve diagnoses.

1.7 Secure E-Voting

Supervisor(s): Helger Lipmaa

Estonia is one of the leading countries in the world with applications like e-voting (and other applications related to e-government). All e-applications must continue correct functioning after inside or outside attacks. An unavoidable part of correct functioning is the privacy of the data. One of the questions related to privacy is the location privacy that guarantees that data sent by participants (e.g., e-voting ballots) cannot be linked to the data origin. Mix-net is a standard cryptographic tool to obtain location privacy. A mix-net takes as an input a tuple of ciphertexts and outputs another tuple of permuted and rerandomized ciphertexts. Such a mixing is accompanied by a zero-knowledge proof that it was done correctly. The University of Tartu participates in an H2020 grant PANORAMIX, the goal of which is to develop a mix-net infrastructure for Europe. The goal of this PhD thesis is to study the security of e-voting systems from a more global viewpoint to find out what are the actual security requirements for mix-nets, and how are security definitions of different components related to each other.

1.8 Ensemble learning under context changes

Supervisor(s): Meelis Kull

Test automation is a widely-used approach to reduce the cost and increase the effectiveness of software testing. However, if test automation is not planned and conducted properly, it will not develop its full potential. Deciding what test activity to automate and what parts of a given System Under Test (SUT) should be tested automatically, at which level of intensity, and with what technique remains a challenge for test engineers and researchers alike. This thesis project aims at developing a customizable test automation analysis framework which can be used as a recommender system with regards to what, how, and how much to automate along the software testing process in a given development context.
Recommendations will be based on data retrieved from software repositories. Machine learning techniques will be applied to generate recommendations. The final results of the project will include a prototypical tool and a comprehensive evaluation of the developed method and tool.