Bachelors projects – Software Engineering Research Group

Notes:

- If you are interested in one of the topics below, please contact the corresponding supervisor, as indicated below each topic title.
- For an up-to-date list of topics, check our web site:
  http://sep.cs.ut.ee/Main/StudentProjects2015#Bachelors

Rescue event categorisation

Supervisor: Siim Karus (siim04 ät ut.ee)

In this thesis, you will analyze data provide by the Rescue services in order to find commonalities in rescue events so as to categorise them. One of the aims will be to isolate and characterize less common rescue event categories, which are of special interest to the Rescue services.

The thesis will be conducted in cooperation with the Rescue Services. The thesis can be written in Estonian.

Workflow Automation With Business Data Streams

Supervisor: Peep Küngas (peep.kungas ät ut.ee)

There exist services such as Flowxo and IFTTT which facilitate automation of simple workflows by facilitating creation of trigger / action pairs or if-then recipes, whose execution will orchestrate applications wrt external stimuli (e.g. application data stream). An example of a popular recipe is the following: "IF I post a picture on Instagram THEN save the photo to Dropbox" or a more complex example is "for every new deal in a CRM send the deal with an e-mail to a person with the GMail service, then wait for about 1 day and send a reminder SMS via Twilio service.

Such systems mostly rely on proprietary application data while there are cases where external stimulus will provide extra benefits. An example of such a case is integration of CRM and credit management tools with external stimuli in form of streaming company debt and risk score data for Order-to-Cash business process. There is a Stream API for business data currently under development at Register OÜ and it will provide a stream of events such as company debt change, changes in board membership and data about newly registered companies. Such data changes events can be easily applied in the context of CRM and a credit management (CM).

The aim of the project will be to leverage provision of an analogue of IFTTT, where users can define recipes for reacting into business data changes via actions in applications such as GMail, Odoo CRM etc.

The project will be done in collaboration with Register OÜ. The application will be developed by using the Complex Event processing (CEP) feature of Register Stream API.

Lead generator for accelerating B2B sales

Supervisor: Peep Küngas (peep.kungas ät ut.ee)

Companies care a lot about improving their sales and to meet such a demand numerous online solutions have been proposed. While for B2C sales the prevalent solutions use social media campaign and Web visitor data, for B2B sales there are solutions, which allow generating a list of leads based on a set of attribute values over company data such as its activity field, size, financial metrics etc.

Some solutions for the Estonian market include https://www.baltictarget.eu, http://sihtgrupid.ee, http://turundusnimekirjad.ee/ and http://www.kliendibaas.ee/. However, these solutions have the following deficiencies:

1. The market segment must be known before generating the leads
2. The set of attributes is mostly limited to geographic, activity field and financial data
3. the data is returned as a file.

This project aims at innovating the B2B sales by providing a solution, which differs from the existing ones in the following way:

1. instead of a list of feature / value pairs a user can define its market segment by giving a set of prospective clients as input to lead generation;
2. in addition to the activity fields, company size and financial metrics also data about owned real estate, credit history, credit risk, media coverage and related persons can be used;
3. instead of outputting leads to a CSV file, lead data will be directly imported to an existing CRM system or a new cloud instance of a CRM will be deployed and populated with the leads.

The project will be done in collaboration with Register OÜ.

Lightning-Fast Multi-Level SOAP-JSON Caching Proxy

Supervisor: Peep Küngas (peep.kungas ät ut.ee)

In a previous Master thesis a solution was developed for proxying SOAP requests/responses to JavaScript widgets exchanging messages with JSON payload. Although this approach was shown to be useful for surfacing Deep Web data, it suffers from some performance bottlenecks, which arise when a SOAP endpoint is frequently used.

This Bachelors thesis aims at developing a cache component, which will make dynamic creation of SOAP-JSON proxies more effective with respect to runtime latency. The resulting cache component will be evaluated from the performance point of view.

A Crawler for RESTful, SOAP Services and Web Forms

Supervisor: Peep Küngas (peep.kungas ät ut.ee)

The Deep Web, consisting of online databases hidden behind SOAP-based or REST-ful Web services or Web forms, is estimated to contain about 500 times more data than the (visible) Web. Despite many advances in search technology, the full potential of the Deep Web has been left largely underexploited. This is partially due to the lack of effective solutions for surfacing and visualizing the data. The Deep Web research initiative at University of Tartu's Institute of Computer Science has developed an experimental platform to surface and visualize Deep Web data sources hidden behind SOAP Web service endpoints. However, currently this experimental platform only supports a limited set of SOAP endpoints, updated on ad hoc basis.

The aim of this project is to build a crawler and an indexing engine capable of recognizing endpoints behind Web forms, RESTful services and SOAP-based services, together with their explicit descriptions (e.g. WSDL interface descriptions, when available). Furthermore, the crawler should

identify examples of queries that can be forwarded to those endpoints, especially for endpoints with no explicit interface descriptions such as Web forms.

This project is available both for Master and for Bachelor students. The goal of the Masters project would be to build a crawler supporting endpoints with and without explicit interfaces. The goal of the Bachelor thesis will be to crawl WSDL interfaces only.

Reverse engineering the RESTBucks's API

Supervisor: Luciano García-Bañuelos (luciano dot garcia ät ut dot ee)

In resent years, multiple tools have emerged that allow one to produce interactive documentation for RESTful applications, often referred to as API blueprints. An API blueprint is basically a Web application that describes the set of resources and operations on them and, on the other hand, provides a way to test the functionality of the application.

The goal of this project is to take an open-source RESTful application, to reverse engineer its API and to specify it using two tools, namely Apiary's Blueprint and Swagger. The project will allow you to critically compare the two tools.

Web Front-End for BIMP

Supervisor: Luciano García-Bañuelos (luciano dot garcia ät ut dot ee)

Business process simulation is a valuable tool for understanding the trade-offs when (re-)designing business processes. For instance, it provides ways for assessing the impact of changes on business processes, thus providing valuable insights to business analysts when it comes to decide how to proceed in the implementation of changes.

A few years ago, our group developed a simulation tool, called [BIMP](#). BIMP is available as "Software as a Service" and is currently used by several Universities in their courses of Business Process Management. Currently, BIMP uses a form-based interface to enter the simulation information. The goal of this project is to implement a Javascript front-end application (probably using JointJS) that renders the BPMN model and allows the user to pick BPMN elements to enter the simulation information. For this project, we expect the student to have knowledge of Javascript and proficiency with Java-based web application development.

Web Front-End for BPMN-Miner Tool

Supervisor: Luciano García-Bañuelos (luciano dot garcia ät ut dot ee)

Our research group has developed a Java-based tool (namely BPMN-Miner) that takes as input a business process execution log (extracted from an enterprise system) and generates a business process model captured in the standard [BPMN notation](#) (in XML format).

The goal of this project is to implement a Javascript-based web front-end that will expose the functionality of the BPMN-Miner tool online. The front-end application will allow users to upload logs, and will graphically render the resulting BPMN models on the browser. For this project, we expect the student to have knowledge of Javascript and be familiar with Java-based web application development.

Runtime Conformance Checking of Control and Data Flow

Supervisor: Fabrizio Maggi (f.m.maggi ät ut dot ee)

Runtime Conformance Checking is considered an important building block in the Business Process Management lifecycle for reasons such as timely detection of non-compliance as well as provision of reactive and proactive countermeasures. In particular, compliance monitoring is related to operational decision support, which aims at extending the application of process mining techniques to on-line, running process instances, so as to detect deviations, recommend what to do next and predict what will happen in the future instance execution.

Runtime Conformance Checking with Fuzzy Logic

Supervisor: Fabrizio Maggi (f.m.maggi ät ut dot ee)

In different research fields a research issue has been to establish if the external, observed behavior of an entity is conformant to some rules/specifications/expectations. Most of the available systems, however, provide only simple yes/no answers to the conformance issue. Some works introduce the idea of a gradual conformance, expressed in fuzzy terms. The conformance degree of a process execution is represented through a fuzzy score.

Deviance Mining of Business Processes

Supervisor: Fabrizio Maggi (f.m.maggi ät ut dot ee)

Deviance mining leverages information hidden in business process execution logs in order to provide guidance to stakeholders so that they can steer the process towards consistent and compliant outcomes and higher process performance. Deviance mining deals with the analysis of process execution logs off-line in order to identify typical deviant executions and to characterize deviance that leads to better or to worse performance. This technique enables evidence-based management of business processes, where process workers and analysts continuously receive guidance to achieve more consistent and compliant process outcomes and a higher performance.

Predictive Monitoring of Business Processes

Supervisor: Fabrizio Maggi (f.m.maggi ät ut dot ee)

Modern information systems that support complex business processes generally maintain significant amounts of process execution data, particularly records of events corresponding to the execution of activities (event logs). Predictive monitoring approaches to analyze such event logs in order to predictively monitor business process executions. When an activity is being executed, they can identify input data values that are more (or less) likely to lead to a good outcome of the process.

Generating Synthetic Event Logs for Benchmarking Process Mining Algorithms

Supervisor: Fabrizio Maggi (f.m.maggi ät ut dot ee)

One way to test a process discovery technique is to generate an event log by simulating a process model, and then verify that the process discovered out of such log matches the original one. For this reason, a tool for generating event logs starting from declarative process models becomes vital for the evaluation of declarative process discovery techniques. The aim of this thesis is to implement an approach for the automated generation of event logs, starting from process models that are based on Declare, one of the most used declarative modeling languages in the process mining literature.

Discovering Business Rules from Event Logs

Supervisor: Fabrizio Maggi (f.m.maggi ät ut dot ee)

Process mining techniques can be used to effectively discover process models from logs that capture a sample of business process executions. Cross-correlating a discovered model with information in the log can be used to improve the underlying process. However, the existing process discovery techniques produce models that tend to be large and complex, especially in flexible environments where process executions involve multiple alternatives. This "overload" of information is caused by the fact that traditional discovery techniques construct procedural models explicitly showing all possible behaviors. Using a declarative model, the discovered process behavior is described as a (compact) set of business rules. Three sub-topics can be investigated in this scenario:

- Discovering Business Rules from Event Logs (control flow)
- Discovering Business Rules from Event Logs (data flow)
- Discovering Business Rules from Event Logs (activity lifecycles)

Plugin for Discovering Business Rules from Event Logs

Supervisor: Fabrizio Maggi (f.m.maggi ät ut dot ee)

The discovery of business rules from event logs is emerging as a new challenge in the business process management field. The actual process behavior as recorded in execution traces is described as a set of business rules (expressed, e.g., using linear temporal logic) and used for process analysis. The candidate is required to implement a (Java) plug-in for a well-known process mining framework called ProM. The code for discovering business rules is already available in C code, so the main task is to create a wrapper for this code and integrate it with ProM.

In this project, we will not assume that you have any prior knowledge in the field of business process management or process mining. We will give you all the input knowledge in this field that you will require to complete the project. All you need are Bachelors-level (Java) programming skills, basic knowledge of XML, and a strong desire to learn new things.

Critical Comparison of the Business Motivation Model (BMM) and i*

Raimundas Matulevicius (raimundas.matulevicius ät ut.ee)

The Business Motivation Model (BMM) is a standardized modeling language to capture important concepts about why a business is undertaking certain actions, such as developing an information system. On the other hand, i* is a modeling language to specify actors and their goals, during the early phases of information system development. There are clear overlaps between these two languages, but also some relevant differences.

The questions to be answered in this thesis are: How does BMM compares against i*? Do they address essentially the same perspective? Do they complement each other? This question will be approached by defining a correspondence between the concepts in these languages and by applying them to a concrete case study.

Comparison of BPMN Security Extensions

Raimundas Matulevicius (raimundas.matulevicius ät ut.ee)

Recently a lot of BPMN extensions are proposed towards security analysis. These extensions concern different aspects, starting from the security problem definition, to security requirements introduction and control identification. The goal of the thesis is develop systematic and coherent overview of the

extensions and to define a set of guidelines for selecting certain BPMN security extensions for targeted problems. This should provide an overview of emerging trends.

Starting points:

1. R. Braun, W. Esswein, Classification of Domain-Specific BPMN Extensions, The Practice of Enterprise Modeling Lecture Notes in Business Information Processing Volume 197, 2014, pp 42-57
2. Menzel, M., Thomas, I., Meinel, C.: Security Requirements Specification in Service- oriented Business Process Management. In: ARES 2009, pp. 41–49 (2009)
3. Altuhhova, O., Matulevičius, R., Ahmed, N.: An extension of business process model and notation for security risk management. International Journal of Information System Modeling and Design (IJISMD) 4(4), 93–113 (2013)
4. Cherdantseva Y., Hilton J., Rana O., Towards SecureBPMN - Aligning BPMN with the Information Assurance and Security Domain, Business Process Model and Notation, Lecture Notes in Business Information Processing Volume 125, 2012, pp 107-115
5. Marcinkowski, B., Kuciapski, M.: A business process modeling notation extension for risk handling. In: Cortesi, A., Chaki, N., Saeed, K., Wierzchoń, S. (eds.) CISIM 2012. LNCS, vol. 7564, pp. 374–381. Springer, Heidelberg (2012)
6. Saleem, M., Jaafar, J., Hassan, M.: A domain-specific language for modelling security objectives in a business process models of soa applications. AISS 4(1), 353–362 (2012)
7. Rodriguez, A., Fernandez-Medina, E., Piattini, M.: A bpmn extension for the modeling of security requirements in business processes. IEICE Transactions on Information and Systems 90(4), 745–752 (2007)

Security Requirements Prioritisation

Raimundas Matulevicius (raimundas.matulevicius ät ut.ee)

Requirements prioritisation plays an important role in the software system development. But, is prioritisation of security requirements in particular different from the traditional prioritisation of requirements? In this thesis it is important to give an answer what prioritisation method and under what circumstance could be applied for security requirements prioritisation. The potential contribution of this thesis is a proposal of the security requirements prioritisation method and its empirical validation.

Starting points:

1. Kaur G., Bawa S., A Survey of Requirement Prioritization Methods, International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue 5, May – 2013
2. Achimugu P., Selamat A., Ibrahim R., Mahrin M. N., A systematic literature review of software requirements prioritization research, Information and Software Technology, Volume 56 Issue 6, June, 2014, 568-585
3. Massey A. K., Otto P. N., Anto?n A. I., Prioritizing Legal Requirements, 2009 Second International Workshop on Requirements Engineering and Law (relaw'09)
4. Park K.-Y., Yoo S.-G., Kim J., Security Requirements Prioritization Based on Threat Modeling and Valuation Graph, Convergence and Hybrid Information Technology Communications in Computer and Information Science Volume 206, 2011, pp 142-152

Lab Package Development & Evaluation for the Course 'Software Testing' (MTAT.03.159)

Supervisor: Dietmar Pfahl (firstname dot lastname at ut dot ee)

The course Software Testing (MTAT.03.159) has currently 6 labs (practice sessions) in which 2nd and 3rd year BSc students learn a specific test technique. We would like to improve existing labs and add new labs.

This topic is intended for students who have already taken this software testing course and who feel that they can contribute to improving it and by the same token complete their Bachelors project. The scope of the project can be negotiated with the supervisor to fit the size of a Bachelors project.

The tasks to do for this project are as follows:

- Selection of a test topic for which a lab package should be developed (see list below)
- Development of the learning scenario (i.e., what shall students learn, what will they do in the lab, what results shall they produce, etc.)
- Development of the materials for the students to use
- Development of example solutions (for the lab supervisors)
- Development of a grading scheme
- Evaluation of the lab package

Topics for which lab packages should be developed (in order of urgency / list can be extended based on student suggestions):

- Web-based Testing with Selenium
- Mutation Testing
- Unit Testing / TDD
- Static Code Analysis
- Search-Based Testing