Doctoral Research Days at FIT 2020

November 20 and November 27, 2020

FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE

Abstract Proceedings.

Doctoral Research Days at FIT 2020: Program and contents

1 Friday 20 November 2020

Session 1

13:00 — Radovan Červený: Generating faster algorithms for d-Path Vertex Cover . . 5

13:15 — Václav Blažej: On the Edge-Length Ratio of 2-Trees 6

13:30 — Matyáš Křišťan: On the m-eternal Domination Number of Cactus Graphs 7

13:45 — Tomáš Pecka: Conversion of Finite Tree Automata to Regular Tree Expressions By State Elimination 8

Coffee break

14:15 — Jan Blizničenko: Type inference for generating UML from Smalltalk 9

14:30 — Marek Skotnica: Design of Systems Supporting Compliance Management 10

14:45 — Jan Slifka: Evolvable Architecture of Client Applications with the use of Normalised Systems Theory 11

15:00 — Vojtěch Knaisl: Designing and implementing machine-actionability for improving evolvability in datastewardship planning in accordance with the Normalised Systems Theory 12

Coffee break

15:30 — Marek Suchánek: Towards Evolvable Architecture of a Normalized Systems Gateway Ontology for Conceptual Models 13

15:45 — David Šenkýř: Processing, Checking, and Modelling of Textual Requirements Specifications 15

16:00 — David Bernhauer: Analysing Indexability of Intrinsically High-Dimensional Data Using TriGen 16

Coffee break

16:30 — Tomáš Beneš: Optimization of high-speed networking applications using hardware architectures 17

16:45 — Karel Hynek: Encrypted traffic monitoring using side channel information 19

17:00 — Jiří Khun: Classification of HPC tasks, communication aspects and their influence on performance 20

2 Friday 27 November 2020

Session 2

13:00 — Štěpán Plachý: Run of incomplete k-local deterministic finite automata in sublinear time 21

13:15 — Jan Ječmen: Contextual Dispatch for Function Specialization 22

13:30 — Filip Kodýtek: Physical Unclonable Functions on FPGAs 23

13:45 — Claudio Kozický: Joint Direct and Transposed Sparse Matrix-Vector Multiplication for Multithreaded CPUs 24

Coffee break

14:15 — Petr Pulc: Advances in Information Retrieval from Video Signal 25

14:30 — Tomáš Nováček: User interface for virtual reality 26

14:45 — Radomír Žemlička: Sequential Bayesian Modeling of Discrete Random Variables 27

15:00 — Jakub Žitný: Deep Generative Models and the explainability of neural networks 29

Coffee break

15:30 — Stanislav Jeřábek: Dummy Rounds Method as Countermeasure against Side Channel Attacks 30
15:45 — Jan Řezníček: Calculating of Non-Homogeneous Continuous Time Markov Chains 32
16:00 — Petr Socha: Side-Channel Analysis and Effective Implementations of Embedded Cryptography 33
16:15 — Tomáš Kolárik: SAT Modulo Differential Equation Simulations 35

Friday 20 November 2020

Session 1

Session chair
Pavel Tvrdík

Doctoral Research Days at FIT 2020

Generating faster algorithms for d-Path Vertex Cover

Radovan Červený

The problem of $\mathcal{F}$-Subgraph Vertex Deletion, $\mathcal{F}$-SVD lies in determining a subset S of vertices of a given graph G such that no subgraph of $G \setminus S$ is isomorphic to some graph in the forbidden set of finite connected graphs $\mathcal{F}$. As this problem is a special case of d-Hitting Set, algorithms faster than trivial enumeration are known for the $\mathcal{F}$-SVD problem.

We aim to tackle problems that can be expressed as the $\mathcal{F}$-SVD problem and for that we propose a general framework for automated design of branching algorithms that solve said problem. Famous problems like Vertex Cover or d-Path Vertex Cover, d-PVC can be expressed as the $\mathcal{F}$-SVD problem. We currently focus on the d-Path Vertex Cover where the forbidden set $\mathcal{F}$ contains only a path on d vertices. By applying our approach to the d-PVC problem, when the d is small we are able to present new branching algorithms which are faster than the currently known state-of-the-art.

Presented results are yet to be published.

List of publications:

Published:

[I] **Radovan Červený**, Ondřej Suchý; *Faster FPT Algorithm for* 5-Path Vertex Cover

In: Proceedings of the 44th International Symposium on Mathematical Foundations of Computer Science, MFCS 2019, August 26-30, 2019, Aachen Germany. ISBN 978-3-95977-117-7, LIPICS Vol. 138 http://www.dagstuhl.de/dagpub/978-3-95977-117-7

Full version available: http://arxiv.org/abs/1906.09213

Doctoral Research Days at FIT 2020

On the Edge-Length Ratio of 2-Trees

Václav Blažej

We study planar straight-line drawings of graphs that minimize the ratio between the length of the longest and the shortest edge. We answer a question of Lazard et al. [Theor. Comput. Sci. **770** (2019), 88–94] and, for any given constant r, we provide a 2-tree which does not admit a planar straight-line drawing with a ratio bounded by r. When the ratio is restricted to adjacent edges only, we prove that any 2-tree admits a planar straight-line drawing whose edge-length ratio is at most $4+\varepsilon$ for any arbitrarily small $\varepsilon > 0$, hence the upper bound on the local edge-length ratio of partial 2-trees is 4.

Presented results are part of [I] .

List of publications:

Accepted:

[I] **Václav Blažej**, Giuseppe Liotta, Jiří Fiala; *On the Edge-Length Ratio of* 2*-Trees* To be in: Proceedings of the 28th International Symposium on Graph Drawing and Network Visualization.

Submitted:

[II] **Václav Blažej**, Michal Opler, Matas Šileikis, Pavel Valtr; *On the Intersections of Non-homotopic Loops* Submitted to CALDAM 2021.

Doctoral Research Days at FIT 2020

On the m-eternal Domination Number of Cactus Graphs

Jan Matyáš Křišťan

Given a graph G, guards are placed on the vertices of G. Then the vertices are subject to an infinite sequence of attacks. Each attack must be defended by a guard moving to the attacked vertex from a neighboring vertex. The m-eternal domination number is the minimum number of guards such that the graph can be defended indefinitely. In this paper we study the m-eternal domination number of cactus graphs, that is, connected graphs where each edge lies in at most one cycle, and we consider three variants of the m-eternal domination number: first variant allows multiple guards to occupy a single vertex, second variant does not allow it, and in the third variant additional "eviction" attacks must be defended. We provide a new upper bound for the m-eternal domination number of cactus graphs, and for a subclass of cactus graphs called Christmas cactus graphs, where each vertex lies in at most two biconnected components, we prove that these three numbers are always equal. Moreover, we present a linear-time algorithm for computing them.

Presented results are part of [I] .

List of publications:

Published:

[I] Václav Blažej, **Jan Matyáš Křišťan**, Tomáš Valla; *On the m-eternal Domination Number of Cactus Graphs.* In: Filiot E., Jungers R., Potapov I. (eds) Reachability Problems. RP 2019. Lecture Notes in Computer Science, vol 11674. Springer, Cham. `https://doi.org/10.1007/978-3-030-30806-3_4`

Doctoral Research Days at FIT 2020

Conversion of Finite Tree Automata to Regular Tree Expressions By State Elimination

Tomáš Pecka

Regular tree languages can be accepted and described by finite tree automata and regular tree expressions, respectively We describe a new algorithm (presented in [2] [III]) that converts a finite tree automaton to an equivalent regular tree expression. Our algorithm is analogous to the well-known state elimination [1] method of the conversion of a string finite automaton to an equivalent string regular expression. We define a generalised finite tree automaton, the transitions of which read the sets of trees described by regular tree expressions. Our algorithm eliminates states of the generalised finite tree automaton, which is analogous to the elimination of states in converting the string finite automaton. This conversion is in the opposite direction than presented recently [3] [I] .

List of publications:

[I] **Tomáš Pecka**, Jan Trávníček, Radomír Polách, Jan Janoušek; *Construction of a Pushdown Automaton Accepting a Postfix Notation of a Tree Language Given by a Regular Tree Expression* In: 7th Symposium on Languages, Applications and Technologies, SLATE 2018, June 21-22, 2018, Guimaraes, Portugal.

[II] Jan Trávníček, **Tomáš Pecka**, Robin Obůrka, Jan Janoušek; *Forward Linearised Tree Pattern Matching Using Tree Border Array* In: Proceedings of the Prague Stringology Conference 2020. Praha: Czech Technical University in Prague, 2020. p. 61–73. ISBN 978-80-01-06749-9.

[III] **Tomáš Pecka**, Jan Trávníček, Jan Janoušek; *Conversion of Finite Tree Automata to Regular Tree Expressions By State Elimination* In: Proceedings of the Prague Stringology Conference 2020. Praha: Czech Technical University in Prague, 2020. p. 11–22. ISBN 978-80-01-06749-9.

References

[1] John E. Hopcroft, Rajeev Motwani, and Jeffrey D. Ullman. *Introduction to automata theory, languages, and computation - international edition, 2nd Edition*. Addison-Wesley, 2003. ISBN: 978-0-321-21029-6.

[2] Tomáš Pecka, Jan Trávníček, and Jan Janoušek. “Conversion of Finite Tree Automata to Regular Tree Expressions By State Elimination”. In: *Proceedings of the Prague Stringology Conference 2020*. Ed. by Jan Holub and Jan Žďárek. 2020, pp. 11–22.

[3] Tomáš Pecka et al. “Construction of a Pushdown Automaton Accepting a Postfix Notation of a Tree Language Given by a Regular Tree Expression”. In: *7th Symposium on Languages, Applications and Technologies, SLATE 2018, June 21-22, 2018, Guimaraes, Portugal*. Ed. by Pedro Rangel Henriques et al. Vol. 62. OASICS. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018, 6:1–6:12. DOI: 10.4230/OASIcs.SLATE.2018.6. URL: https://doi.org/10.4230/OASIcs.SLATE.2018.6.

Doctoral Research Days at FIT 2020

Type inference for generating UML from Smalltalk

Jan Blizničenko

Generating programming code out of models is a useful approach towards maintainable software systems, yet legacy systems have often no models at all. There are multiple ways to partially automate or aid the process of generating models out of legacy code, yet there is one kind of object-oriented programming languages substantially harder to transform – dynamically typed languages. While statically typed languages enforce programmers to explicitly state the data types of various elements, dynamically typed languages do not, presenting significant difficulties when gathering associations between classes. This paper presents an ongoing effort towards dealing with gathering associations by combining various type inference techniques and tools and how the authors aim to use UML models as a transition form between origin and destination programming languages, producing UML models as useful byproducts. Pharo – Smalltalk-based dynamically typed programming language is used as a case study.

List of publications:

Accepted:

[I] **Jan Blizničenko**, Robert Pergl; *Generating UML Models from Pharo Code* IWST19 — International Workshop on Smalltalk Technologies, 2019. (presented, not yet published).

[II] **Jan Blizničenko**, Robert Pergl; *Towards Generating Software Systems From Legacy Code: Gathering Associations From Smalltalk Software*; 10th Enterprise Engineering Working Conference (EEWC), 2020. (presented, not yet published).

Doctoral Research Days at FIT 2020

Design of Systems Supporting Compliance Management

Marek Skotnica

Compliance management is a process which ensures that a set of people follows a given set of rules. These rules are usually a mix of procedures, documentation, policies, industry standards, and legal obligations. A primary goal of this research is to describe how to design systems which support this domain.

Our approach reviews all steps in building compliance management systems. These usually involve 1) Domain Modeling - a creation of high-quality machine-readable domain models, 2) System Specification - a machine-readable system formalization, 3) System Execution - execution of a formalized system specification, 4) System Optimization - measuring and suggestion of domain model improvements. We do introduce novel approaches in system specification and execution.

We do showcase our methods and techniques on the domain of procedural law and legal contracts. Such systems can be executed as software artefacts or as a Blockchain smart contracts (SC). The SC has enormous potential in automating traditional paper contracts and encoding contract logic into program code. Thus, replacing the role of a notary and a central authority. It may dramatically reduce an effort with administration workload and the enforcement of such contracts.

List of publications:

1. Marek Skotnica, Steven J. H. van Kervel, and Robert Pergl. “Towards the Ontological Foundations for the Software Executable DEMO Action and Fact Models”. en. In: Advances in Enterprise Engineering X. Funchal, Madeira: Springer International Publishing, May 2016, pp. 151–165.

2. Marek Skotnica, Steven J. H. van Kervel, and Robert Pergl. “A DEMO Machine - A Formal Foundation for Executionof DEMO Models”. In: Advances in Enterprise Engineering XI. Ed. by David Aveiro et al. Cham: Springer International Publishing, 2017, pp. 18–32. ISBN: 978-3-319-57955-9.

3. Ondřej Mráz, Robert Pergl, Pavel Náplava, and Marek Skotnica. “Converting DEMO PSI Transaction Pattern into BPMN: A Complete Method”. In: Advancesin Enterprise Engineering XI: 7th Enterprise Engineering Working Conference, EEWC 2017, Antwerp, Belgium, May 8-12, 2017, Proceedings. Ed. by David Aveiro et al. Cham: Springer International Publishing, 2017, pp. 85–98. ISBN: 978-3-319-57955-9

4. Barbora Hornáčková, Marek Skotnica, and Robert Pergl. “Exploring a Role of Blockchain Smart Contracts in Enterprise Engineering”. In: Advances in Enterprise Engineering XII. Ed. by David Aveiro et al. Cham: Springer International Publishing, 2019, pp. 113–127. ISBN: 978-3-030-06097-8.

5. Skotnica M., Pergl R. (2020) Das Contract - A Visual Domain Specific Language for Modeling Blockchain Smart Contracts. In: Aveiro D., Guizzardi G., Borbinha J. (eds) Advances in Enterprise Engineering XIII. EEWC 2019. Lecture Notes in Business Information Processing, vol 374. Springer, Cham. pp. 149-166. ISBN: 978-3-030-37932-2

Doctoral Research Days at FIT 2020

Evolvable Architecture of Client Applications with the use of Normalised Systems Theory

Jan Slifka

Client applications are an essential part of the most software because they provide interfaces with which the users interact. We can think of different dimensions that the client application consist of – a platform, business logic and design system. These dimensions could be independent and reusable to a large extent. However, they usually are tightly coupled, which impede the evolvability.

The problem is also a rapid development of technologies used for client applications (e.g., new frameworks for web applications). Thus these applications become obsolete every few years and have to be rewritten, usually from scratch. Nowadays, with the increasing number of mobile devices, one client application is usually not enough. Each platform needs a specific client application. Therefore, more applications have to be rewritten, which tremendously increases the costs and time demand for software development.

The research of the Normalised Systems at the University of Antwerp focuses on evolvable systems, where the domain and technical parts are separated. So-called expansion is used to generate the code from the domain specification. Systems can be regenerated to the new technologies without the loss of the domain layer.

The goal of this research is to use existing and design new ontologies and modelling languages to describe different dimensions of client applications and use them for generating the actual software using expanders following the theory of Normalised Systems. The main focus is on business applications which usually have a limited set of features with only minor customisations, so they have the highest potential to save time and the costs.

List of publications:

Published:

[I] Suchánek, M.; **Slifka, J.** *Evolvable and Machine-Actionable Modular Reports for Service-Oriented Architecture.* In: Enterprise and Organizational Modeling and Simulation. Springer, Cham, 2019. p. 43-59. 1. vol. 366. ISSN 1865-1348. ISBN 978-3-030-35645-3.

[II] Pergl, R.; Hooft, R.; Suchánek, M.; Knaisl, V.; **Slifka, J.** *"Data Stewardship Wizard": A Tool Bringing Together Researchers, Data Stewards, and Data Experts around Data Management Planning.* Data Science Journal. 2019, 18(1), 1-8. ISSN 1683-1470.

[III] **Slifka, J.**; Pergl, R. *Laying the Foundation for Design System Ontology.* In: Trends and Innovations in Information Systems and Technologies. Springer, Cham, 2020. p. 778-787. ISSN 2194-5357. ISBN 978-3-030-45687-0.

Doctoral Research Days at FIT 2020

Designing and implementing machine-actionability for improving evolvability in datastewardship planning in accordance with the Normalised Systems Theory

Ing. Vojtěch Knaisl

My topic is to focus on the problem of evolvability in Data Stewardship Planning. In general, the problem of evolvability and long-term sustainability is a big challenge nowadays. My focus is to improve the evolvability in the Data Stewardship Planning process.

In terms of the increasing amount of data in research experiments, the importance of right planning, storing data, and long-term sustainability increase too. Data Stewardship (DS) contains all these phases. In CCMi [1], we are developing a unique tool Data Stewardship Wizard [2]. In this area, we face a problem with evolvability due to changing requirements and standards on DS in every phase of the projects, further, the changing requirements on output and format of desired plans.

During my first two years, I focused on the general evolvability of structured and well-formatted documents. I wrote two articles, participated in one journal paper, and made some proposals. But because the topic was too wide, we narrow it to the Data Stewardship domain, where I continue on my research. My goal to apply Normalized Systems Theory [3] and other formal methods for creating a methodical framework and proposal of a technical solution for Data Stewardship Planning. The Normalized Systems Theory developed at the University of Antwerp brings the principles for the construction of evolvable systems. This general theory was applied, among other things, on the evolvability of documents [4].

List of publications:

Accepted:

[I] **Vojtěch Knaisl**; *Proposing an Architecture of an Intelligent Evolvable Document Generation System based on the Normalized Systems Theory*: Proceedings of the 15$^{\text{th}}$ International Workshop on Enterprise & Organizational Modeling and Simulation, pp. 70-81, Rome, Italy, 2019. (EOMAS)

[II] **Vojtěch Knaisl**, Robert Pergl, Jan Slifka, Marek Suchánek, Rob Hooft; *"Data Stewardship Wizard": A Tool Bringing Together Researchers, Data Stewards, and Data Experts around Data Management Planning*: Codata Science Journal. 2019, 18(1), 1-8. ISSN 1683-1470. (EOMAS)

[III] **Vojtěch Knaisl**, Robert Pergl; *Proposing Ontology-Driven Content Modularization in Documents Based on the Normalized Systems Theory*: WorldCist'20 - 8$^{\text{th}}$ World Conference on Information Systems and Technologies, pp. 45-54, Budva, Montenegro, 2020.

References

[1] *Centre for Conceptual Modelling [online], [cit. 2019-10-14].* URL: https://ccmi.fit.cvut.cz.

[2] *Data Stewardship Wizard [online], [cit. 2019-10-14].* URL: https://ds-wizard.org.

[3] Herwig Mannaert, Jan Verelst, and Peter De Bruyn. *Normalized Systems Theory: From Foundations for Evolvable Software toward a General Theory for Evolvable Design.* 2016. ISBN: 978-90-77160-09-1.

[4] Gilles Oorts, Herwig Mannaert, and Peter De Bruyn. "Exploring Design Aspects of Modular and Evolvable Document Management". In: *Advances in Enterprise Engineering XI.* Ed. by David Aveiro et al. Cham: Springer International Publishing, 2017, pp. 126–140. ISBN: 978-3-319-57955-9.

Doctoral Research Days at FIT 2020

Towards Evolvable Architecture of a Normalized Systems Gateway Ontology for Conceptual Models

Marek Suchánek

Conceptual modelling as a crucial part of software engineering and other domains is used to precisely describe a domain of interest in order to promote understanding and communication between people. A conceptual model is used mainly in the analysis phase of software development and for requirements specification. However, Model-Driven Development deals with re-use of conceptual models to generate a software system or at least its generic parts, or so-called skeletons and fragments. These methods usually suffer from the detachment of a model and generated software when changes are incorporated. The evolvability issues are targeted by the Normalized Systems (NS), and it does so using its models composed of so-called Elements. The NS theory is widely applicable as we demonstrated for documents [1] [2] or for messages in Service-Oriented Architecture [6] . Our research targets the ultimate goal to allow bi-directional and seamless transformation between models of Normalized System and conceptual models in various modelling languages, such as UML, OntoUML, ORM, or BPMN. Resource Description Framework (RDF) and Web Ontology Language (OWL) technologies provide versatility that can be used to maintain evolvability and extensibility of the solution. The modular architecture of this ontology-based solution further leverages evolvability and also separates its development to specific parts. The transformation between NS and OWL has been successfully implemented as prototype [8] . The prototype has been then refined into an NS application itself, i.e., to be expandable using specific version NS metamodel [IX] . The NS metamodel is meta-circular, which allows the tool to be used practically for any other NS model. Bridging the gap between NS models and conceptual models using a gateway ontology allows to integrate models in both directions and even combine knowledge from multiple conceptual models. Additionally, during the mapping conceptual modelling languages, enhancements for the NS metamodel are expected to be proposed, for example, related to inheritance implementation patterns [3] . Another interesting aspect is related to mapping conceptual modelling languages between themselves to allow their seamless combination and semantic integration for both structural [4] and behavioural [5] models. The next step is to proceed with the mapping of UML as the most widely-used language in software analysis and design. A review of possible approaches to transform UML models to OWL and vice versa is already provided in [7] .

List of publications:

Published:

[1] **Marek Suchánek**, Robert Pergl; *Evolvable Documents – an Initial Conceptualization.* In: PATTERNS 2018, The Tenth International Conference on Pervasive Patterns and Applications. Wilmington: IARIA, 2018. p. 39-44. ISSN 2308-3557. ISBN 978-1-61208-612-5.

[2] **Marek Suchánek**, Robert Pergl; *Towards Evolvable Documents with a Conceptualization-Based Case Study.* International Journal on Advances in Intelligent Systems. 2018, 11(3&4), 212-223. ISSN 1942-2679.

[3] **Marek Suchánek**, Robert Pergl; *Evolvability Evaluation of Conceptual-Level Inheritance Implementation Patterns.* In: PATTERNS 2019, The Eleventh International Conference on Pervasive Patterns and Applications. Wilmington: IARIA, 2019. p. 1-6. ISSN 2308-3557. ISBN 978-1-61208-612-5.

[4] **Marek Suchánek**, Robert Pergl; *Designing an Ontology for Semantic Integration of Various Conceptual Models.* In: EOMAS 2019, 15th International Workshop on Enterprise and Organizational Modeling and Simulation. Springer, Cham, 2019. p. 3-17. 1. vol. 366. ISSN 1865-1348. ISBN 978-3-030-35645-3.

[5] **Marek Suchánek**, Robert Pergl; *Mapping UFO-B to BPMN, BORM, and UML Activity Diagram* In: EOMAS 2019, 15th International Workshop on Enterprise and Organizational Modeling and

Simulation. Springer, Cham, 2019. p. 82-98. 1. vol. 366. ISSN 1865-1348. ISBN 978-3-030-35645-3.

[6] **Marek Suchánek**, Jan Slifka, Robert Pergl; *Evolvable and Machine-Actionable Modular Reports for Service-Oriented Architecture.* In: EOMAS 2019, 15th International Workshop on Enterprise and Organizational Modeling and Simulation. Springer, Cham, 2019. p. 43-59. 1. vol. 366. ISSN 1865-1348. ISBN 978-3-030-35645-3.

[7] **Marek Suchánek**, Robert Pergl; *Case-Study-Based Review of Approaches for Transforming UML Class Diagrams to OWL and Vice Versa.* In: 2020 IEEE 22nd Conference on Business Informatics (CBI). Los Alamitos: IEEE Computer Society, 2020. p. 270-279. vol. 1. ISBN 978-1-7281-9926-9.

[8] **Marek Suchánek**, Herwig Mannaert, Peter Uhnák, Robert Pergl; *Bi-directional Transformation between Normalized Systems Elements and Domain Ontologies in OWL.* In: Proceedings of the 15th International Conference on Evaluation of Novel Approaches to Software Engineering. Porto: SciTePress - Science and Technology Publications, 2020. p. 74-85. ISSN 2184-4895. ISBN 978-989-758-421-3.

Submitted:

[IX] **Marek Suchánek**, Herwig Mannaert, Peter Uhnák, Robert Pergl; *Towards Evolvable Ontology-Driven Development with Normalized Systems.* In: Communications in Computer and Information Science, Springer, 2020.

Doctoral Research Days at FIT 2020

Processing, Checking, and Modelling of Textual Requirements Specifications

David Šenkýř

The quality of Requirements Engineering plays an essential role in the life cycle of every project – because the other phases depend on it. Writing requirements specifications in natural language is a common practice. Unfortunately, the natural language is prone to several inaccuracies. We already tackled the problems of *ambiguity* [II] and *incompleteness* [III] [V] . We now proceed with the problem of *incompleteness*.

Our approach is to investigate methods of *grammatical inspection* to identify patterns in requirements specification written in the textual form [I] . Based on the patterns, semantic networks (e.g. *ConceptNet* and *BabelNet*), and pre-configured data from on-line dictionaries, we can extract the information from the text. The result of our work is the TEMOS tool. It helps to eliminate the mentioned problems, and it also generates fragments of models in the form of UML class diagram or SHACL Shapes [IV] .

In the past year, we cooperate with Marek Skotnica on the project of automatic identification of acts via *Organization Essence Revealing* (OER) method that is used in DEMO (Design & Engineering Methodology for Organizations) methodology. For this purpose, we develop the TEMOS OER version of our tool. In this project, we also reuse the approach of the grammatical inspection methods.

List of publications:

Published:

[I] **David Šenkýř**, Petr Kroha; *Patterns in Textual Requirements Specification* In: Proceedings of the 13th International Conference on Software Technologies, pp. 197–204, Porto, Portugal, 2018. SCITEPRESS – Science and Technology Publications. ISBN 978-989-758-320-9.

[II] **David Šenkýř**, Petr Kroha; *Patterns of Ambiguity in Textual Requirements Specification*; In: New Knowledge in Information Systems and Technologies, volume 1, pp. 886–895, Cham, 2019. Springer International Publishing. ISBN 978-3-030-16181-1.

[III] **David Šenkýř**, Petr Kroha; *Problem of Incompleteness in Textual Requirements Specification*; In: Proceedings of the 14th International Conference on Software Technologies, pp. 323–330, Porto, Portugal, 2019. SCITEPRESS – Science and Technology Publications. ISBN 978-989-758-379-7.

[IV] **David Šenkýř**; *SHACL Shapes Generation from Textual Documents* In: Enterprise and Organizational Modeling and Simulation. Springer, Cham, 2019. p. 121–130. 1. vol. 366. ISSN 1865-1348. ISBN 978-3-030-35645-3.

[V] **David Šenkýř**, Petr Kroha; *Patterns for Checking Incompleteness of Scenarios in Textual Requirements Specification* In: Proceedings of the 15th International Conference on Evaluation of Novel Approaches to Software Engineering. Porto: SciTePress - Science and Technology Publications, 2020. p. 289-296. ISSN 2184-4895. ISBN 978-989-758-421-3.

Doctoral Research Days at FIT 2020

Analysing Indexability of Intrinsically High-Dimensional Data Using TriGen

David Bernhauer

The TriGen algorithm is a general approach to transform distance spaces in order to provide both exact and approximate similarity search in metric and non-metric spaces. This paper focuses on the reduction of intrinsic dimensionality using TriGen. Besides the well-known intrinsic dimensionality based on distance distribution, we inspect properties of triangles used in metric indexing (the triangularity) as well as properties of quadrilaterals used in ptolemaic indexing (the ptolemaicity). We also show how LAESA with triangle and ptolemaic filtering behaves on several datasets with respect to the proposed indicators.
Presented results are part of [I] .

List of publications:

Published:

[I] **David Bernhauer**, Tomáš Skopal; *Analysing Indexability of Intrinsically High-Dimensional Data Using TriGen* In: Similarity Search and Applications. Springer, pp. 261–269, 2020. ISBN 978-3-030-60936-8.

Doctoral Research Days at FIT 2020

Optimization of high-speed networking applications using hardware architectures

Tomáš Beneš

Modern society is becoming more and more dependent on the interconnectivity of computer systems. Nowadays, almost everybody using several of these interconnected systems on an everyday basis. This produces an enormous quantity of data that needs to be transferred across multiple systems, usually located in different countries. High-speed networks, which enable such transmission, have become the backbone of today's society. For a long time, a specialized ASIC hardware implementation dominated these systems. However, their narrow specialization and inflexibility to adapt to new applications are the leading causes of their gradual replacement by the FPGA systems. Specialized implementations inside FPGA have more flexibility than ASICs with comparable performance. Therefore, it outclasses the general-purpose CPUs by orders of magnitude. These implementations are becoming widespread mostly in cloud-based solutions, where a single accelerator card or network card incorporating FPGA can significantly increase performance and lower the maintenance cost.

However, high-speed processing the network traffic reveals many research challenges and difficulty of the hardware design rapidly grows with the increasing network speed. Even though modern technology provides more resources and allows for reaching higher frequencies of the design, systems that should work at above 100 Gb/s require completely new architectures to fit on chip and get desired performance.

We have already elaborated on various use-cases of hardware accelerated network traffic processing. One of the initial problems we have tackled is designing a low-latency compression unit for the LZ4 scheme aimed for use in 10G applications. As part of this architecture, we have designed a high-throughput match search unit for any lossless compression. The unit is using a novel approach of shared-dictionary between compression cores instead of the traditional multi-dictionary one.

Naturally, our current attention focuses on significantly higher speeds of the network, i.e., above 100 Gb/s traffic. Besides the processing the incoming network packets, our goal is to support and accelerate computation of characteristics of the network traffic, which helps to improve analysis of encrypted traffic. This goal brings many open research challenges, which are unreachable using only the general-purpose CPUs due to their performance limits, especially in parallelism. Our focus is to design a system using a hardware-software co-design paradigm to reach up to 400 Gb/s at link speed without any packet loss.

The first challenge is an insufficient memory capacity inside the chip. A possible solution is to use external memories located on FPGA boards, however, this must be done very effectively. Otherwise, the latency of the external memory subsystem totally disrupts the overall throughput of the device. Therefore, we have addressed this issue by proposing a Pipelined ALU for effective external memory access. Our novel design uses a specialized cache that aggregates data operations and postpones their evaluation until the data are available. Techniques and design challenges in this abstract will be used as a base of the future dissertation thesis.

List of publications:

Published:

[I] M. Bartík, S. Ubik, P. Kubalík, and **T. Beneš**, "Performance Comparison of Multiple Approaches of Status Register for Medium Density Memory Suitable for Implementation of a Lossless Compression Dictionary: (Abstract Only)," in Proceedings of the 2018 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays, New York, NY, USA, Feb. 2018, p. 290, doi: 10.1145/3174243.3174976.

[II] M. Bartík, **T. Beneš**, and P. Kubalík, "Design of a High-Throughput Match Search Unit for Lossless Compression Algorithms," in 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), Jan. 2019, pp. 0732–0738, doi: 10.1109/CCWC.2019.8666521.

[III] **T. Beneš**, M. Bartík, and P. Kubalík, "High Throughput and Low Latency LZ4 Compressor on

FPGA," in 2019 International Conference on ReConFigurable Computing and FPGAs (ReConFig), Dec. 2019, pp. 1–5, doi: 10.1109/ReConFig48160.2019.8994794.

[IV] **T. Beneš**, M. Kekely, K. Hynek, and T. Èejka, "Pipelined ALU for effective external memory access in FPGA," in 2020 23rd Euromicro Conference on Digital System Design (DSD), Aug. 2020, pp. 97–100, doi: 10.1109/DSD51259.2020.00026.

[V] M. Bartík, **T. Beneš**, and P. Kubalík, "An In-Sight Into How Compression Dictionary Architecture Can Affect the Overall Performance in FPGAs," IEEE Access, vol. 8, pp. 183101–183116, 2020, doi: 10.1109/ACCESS.2020.3029691.

[VI] K. Hynek, **T. Beneš**, T. Èejka, and H. Kubátová, "Refined Detection of SSH Brute-Force Attackers Using Machine Learning," in ICT Systems Security and Privacy Protection, vol. 580, M. Hölbl, K. Rannenberg, and T. Welzer, Eds. Cham: Springer International Publishing, 2020, pp. 49–63.

[VII] **T. Beneš**, T. Èejka, and H. Kubátová, "The next step of P4 FPGA architectures: External Memories," in Proceedings of the 8th Prague Embedded Systems Workshop (PESW), 2020.

Doctoral Research Days at FIT 2020

Encrypted traffic monitoring using side channel information

Karel Hynek

Personal privacy on the internet is one of the most discussed topics in recent years. Internet users are becoming more educated and are aware of surveillance and data gathering. Companies are therefore compelled to strengthen services by the use of encryption. The rise of encrypted traffic on the internet is enormous in recent years, generally accepted as a good thing. However, threat actors can also use encryption to hide their malicious activities on the network. Current network monitoring systems are entirely blind in case of encrypted traffic. Our research extends the existing knowledge in IP Flow based network monitoring for traffic classification and threat detection, especially in encrypted traffic. Even though the encrypted content is not available for inspection, the flat traffic shape can disclose a relatively high amount of information. Therefore, we use side-channel features (such as individual packet size, inter-packet time, and packet direction) to create a discriminative feature set. This kind of feature vectors as input data allow us to use Machine Learning, and to design and develop reliable detection and classification algorithms. Using side-channel features, we have already gained many achievements: We have identified novel malicious IoT threats with high precision, increased the accuracy of blacklist based malicious traffic detector, improved detection of SSH brute force attacks, developed recognition algorithm for DNS over HTTPS (DoH) traffic, and classified the DoH clients. Additionally, the side channel information helped us to discover information leakage from the DoH communication by the Firefox Browser, which raises serious doubts about its DoH implementation security. All the mentioned use-cases from network security and privacy area that were covered in this abstract will be used as a base of the future dissertation thesis.

List of publications:

1. Dominik Soukup, Tomáš Čejka, **Karel Hynek**; *Behavior Anomaly Detection in IoT Networks*
   In: Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI - 2019). Cham: Springer International Publishing, 2020. p. 465-473. Lecture Notes on Data Engineering and Communications Technologies. vol. 49. http://dx.doi.org/10.1007%2F978-3-030-43192-1_53

2. **Karel Hynek**, Tomáš Čejka, Martin Žádník, Hana Kubátová; *Evaluating Bad Hosts Using Adaptive Blacklist Filter*
   In: Proceedings of the 9th Mediterranean Conference on Embedded Computing - MECO'2020. Institute of Electrical and Electronics Engineers, Inc., 2020. http://dx.doi.org/10.1109%2FMECO49872.2020.9134244

3. **Karel Hynek**, Tomáš Beneš, Tomáš Čejka, Hana Kubátová; *Refined detection of SSH brute-force attackers using machine learning* In: ICT Systems Security and Privacy Protection. Cham: Springer, 2020. p. 49-63. IFIP Advances in Information and Communication Technology. vol. 580. http://dx.doi.org/10.1007%2F978-3-030-58201-2_4

4. Dmitrii Vekshin, **Karel Hynek**, Tomáš Čejka; *DoH Insight: Detecting DNS over HTTPS by Machine Learning*
   In: Proceedings of the 15th International Conference on Availability, Reliability and SecurityAugust 2020, Article No.: 87, Pages 1–8, https://doi.org/10.1145/3407023.3409192

5. **Karel Hynek**, Tomáš Čejka; *Privacy Illusion: Beware of Unpadded DoH*
   Accepted by IEEE – IEMCON 2020

Doctoral Research Days at FIT 2020

Classification of HPC tasks, communication aspects and their influence on performance

Jiří Khun

High-performance computing is an essential part of computer science, focusing on demanding computations. It is a broad area, integrating varied HW and SW solutions. The ultimate goal is to bring the fastest and effective approaches to solving the most challenging computational tasks. Today's acceleration platforms can vary significantly, but their common feature is a parallel approach to handling the computations. Such accelerators typically consist of many computational elements that divide the tasks, solve individual chunks separately, and possibly (but not necessarily) merge the partial results back into a final solution. That necessarily leads to a design of parallel algorithms that can be handle by the acceleration platforms and maximally utilize their resources. Therefore the HPC tasks have been extensively researched in the past few decades. The research has been focused on generalization and finding common traits among the tasks. The ultimate goal is to create a classification system that would allow the sharing of similar acceleration approaches on appropriate acceleration devices. In other words, the usage of a design pattern convenient for a particular task. Several influential papers related to this topic have been published. They divided the HPC tasks into categories (called families or "dwarves") and recommended methodology for their acceleration. Their primary focus lies in the nature of the computation (sequential/parallel parts of computations), memory accesses, and algorithm internal communication that is often a performance-critical part of the process. The internal communication issue came from the simple fact that the computation is not spread in time as a sequential process but rather in space to do many steps at once in different parts of an accelerator. Our research targets FPGA acceleration that is less explored than other major platforms (CPU, GPU). We have noticed that the current classification of HPC tasks from the FPGA point of view is less precise in terms of the communication aspects, and therefore there is possibly space for improvement. Due to its flexibility, FPGA could offer better performance for some problem-specific or even instance-specific variants of established HPC tasks families, especially in terms of the communication. We try to find such HPC tasks, evaluate their aspects, and improve the classification system accordingly.

List of publications:

1. **Jiří Khun**, Ivan Šimeček; *Parallelization of Artificial Immune Systems Using a massive parallel approach via modern GPUs* In: Proceedings of Seminar Programs and Algorithms of Numerical Mathematics 17. Praha: Matematický ústav AV ČR, 2015. pp. 106-111. ISBN 978-80-85823-64-6.

2. **Jiří Khun**, Ivan Šimeček, Robert Lórencz; *GPU solver for systems of linear equations with infinite precision* In: 17th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing. Los Alamitos: IEEE Computer Society, 2016. pp. 121-124. ISBN 978-1-5090-0461-4.

3. **Jiří Khun**; *Solver for Systems of Linear Equations with Infinite Precision on a GPU Cluster* In: Proceedings of the 20th International Scientific Student Conferenece POSTER 2016. Praha: Czech Technical University in Prague, 2016. ISBN 978-80-01-05950-0.

4. **Jiří Khun**, Martin Novotný, Miroslav Skrbek; *High-Performance Spiking Neural Network Simulator* In: Proceedings of the 8th Mediterranean Conference on Embedded Computing - MECO'2019. Institute of Electrical and Electronics Engineers, Inc., 2019. p. 88-91. ISSN 2377-5475. ISBN 978-1-7281-1739-3.

5. **Jiří Khun**; *Akcelerace imunitních algoritmů pomocí FPGA* In: Sborník příspěvků PAD 2019 - elektronická verze. Praha: AMCA spol. s r.o., 2019. ISBN 978-80-88214-20-5.

Friday 27 November 2020

Session 2

Session chair
Štěpán Starosta

Doctoral Research Days at FIT 2020

Run of incomplete k-local deterministic finite automata in sublinear time

Štěpán Plachý

A synchronizing word for a deterministic finite automaton with an incomplete transition function is a word such that all paths in the automaton's graph labeled with the word end in the same state. A stronger property of k-locality occurs when all words of length at least k are synchronizing. Reading a word with such automaton then either fails or otherwise is either accepted or rejected only based on its suffix of length up to k. Languages accepted by this kind of automata are called stricly locally testable and are characterized by sets of allowed prefixes, forbidden factors and allowed suffixes.

We present an algorithm for a run of incomplete k-local deterministic finite automata using backwards pattern matching technique that in best case can run in sublinear time. The automaton is first analyzed for its characteristic sets of factors which are then matched in an input string using for example Commentz-Walter algorithm. The worst case time complexity for an input string of length n is then $\mathcal{O}(nk)$ while the best case is $\mathcal{O}(n/k)$.

List of publications:

Doctoral Research Days at FIT 2020

Contextual Dispatch for Function Specialization

Jan Ječmen

In order to generate efficient code for dynamic languages, compilers often need information not readily available in the source code. Leveraging a mixture of static and dynamic information, just-in-time compilers can speculate on the missing information. Within one compilation unit, code is specialized to the observed behaviors. We propose an approach to further the specialization, by disentangling classes of behaviors into separate optimization units. With contextual dispatch, functions are versioned and each version is compiled under different assumptions. When a function is invoked, the implementation dispatches to a version that was optimized under assumptions matching the dynamic context of the call. As a proof-of-concept, we describe a compiler for the R language which uses this approach. We evaluate contextual dispatch on a set of benchmarks and compare to traditional speculation with deoptimization techniques. Our implementation is, on average, 1.7× faster than the GNU R reference implementation, and contextual dispatch contributes to the performance significantly in 18 of 46 programs in our benchmark suite.

List of publications:

1. Flückiger, Olivier and Chari, Guido and Yee, Ming-Ho and Ječmen, Jan and Hain, Jakob and Vitek, Jan. 2020. *Contextual Dispatch for Function Specialization* Proc. ACM Program. Lang. 4, OOPSLA, Article 220 (November 2020), 36 pages. https://doi.org/10.1145/3428288

Doctoral Research Days at FIT 2020

Physical Unclonable Functions on FPGAs

Filip Kodýtek

PUFs (Physical Unclonable Function) are increasingly used in proposals of security architectures for device identification and cryptographic key generation. Many PUF designs for FPGAs proposed up to this day are based on ring oscillators (RO). The classical approach is to compare frequencies of ROs and produce a single output bit from each pair of ROs based on the result of comparison of their frequencies. Such ROPUF design requires all ROs to be mutually symmetric and also the number of pairs of ROs is limited in order to preserve the independence of bits in the PUF response. This led us to design a new ROPUF on FPGA which is capable of generating multiple output bits from each pair of ROs and is also allowing to create higher number of pairs of ROs, thereby making the use of ROs more efficient than the classical approach. Our PUF design is based on selecting a particular part of a counter value and using it for the PUF output. In principle, this PUF design does not need the ROs to be mutually symmetric, however, it is shown that this ROPUF design has significantly better properties with varying supply voltage and temperature when symmetric ROs are used. We also compared three similar approaches that are all based on extracting parts of the counter values for the PUF output in terms of the quality of their output. These approaches differ only in the measurement method that is used to obtain the counter values.

Moreover, we observed that we can use the same design as a TRNG (True Random Number Generator). This enables us to use the same design for various applications – PUF can be used for generating and storing cryptographic keys, TRNG for generating session and ephemeral keys, nonces and salts. The proposed TRNG design exhibited satisfactory behaviour as it passed NIST statistical test suite. However, in order to evaluate the TRNG thoroughly, a statistical model of its source of randomness is needed.

List of publications:

1. **Kodýtek, F.**; Lórencz, R.: A design of ring oscillator based PUF on FPGA. In *18th IEEE Symposium on Design and Diagnostics of Electronic Circuits and Systems.* April 22–24, 2015 – Belgrade, Serbia.

2. **Kodýtek, F.**; Lórencz, R.: Proposal and Properties of Ring Oscillator Based PUF on FPGA, 2016. In *Journal of Circuits, Systems and Computers.* March 2016, Vol. 25, No. 03. ISSN 0218-1266.

3. **Kodýtek, F.**; Lórencz, R.; Buček, J.: Improved ring oscillator PUF on FPGA and its properties. In *Microprocessors and Microsystems.* 2016, ISSN 0141-9331, http://dx.doi.org/10.1016/j.micpro.2016.02.005.

4. **Kodýtek, F.**; Lórencz, R.; Buček, J.; Buchovecká, S.: Temperature dependence of ROPUF on FPGA. In *Euromicro Conference on Digital System Design* (Poster). August 31 – September 2, 2016 – Limassol, Cyprus.

5. Buchovecká, S.; Lórencz, R.; **Kodýtek, F.**; Buček, J.: True Random Number Generator based on ROPUF circuit. In *Euromicro Conference on Digital System Design.* August 31 – September 2, 2016 – Limassol, Cyprus

6. Buchovecká, S.; Lórencz, R.; **Kodýtek, F.**; Buček, J.: True random number generator based on ring oscillator PUF circuit. In *Microprocessors and Microsystems.* 2017, ISSN 0141-9331, https://doi.org/10.1016/j.micpro.2017.06.021.

7. Buchovecká, S.; Lórencz, R.; Buček, J.; **Kodýtek, F.**. Lightweight Authentication and Secure Communication Suitable for IoT Devices. In *The International Conference on Information Systems Security and Privacy.* February 25–27, 2020, Malta, pp. 75–83.

8. **Kodýtek, F.**; Lórencz, R.; Buček, J. Comparison of three counter value based ROPUFs on FPGA. In *Euromicro Conference on Digital System Design.* August 26–28, 2020 – Portorož, Slovenia.

Doctoral Research Days at FIT 2020

Joint Direct and Transposed Sparse Matrix-Vector Multiplication for Multithreaded CPUs

Claudio Kozický

Repeatedly performing sparse matrix-vector multiplication (SpMV) followed by transposed sparse matrix-vector multiplication (SpM$^{\mathrm{T}}$V) with the same matrix is a part of several algorithms, for example the Lanczos Biorthogonalisation algorithm and the Biconjugate Gradient Method. Such algorithms can benefit from combining parallel SpMV and SpM$^{\mathrm{T}}$V into a single operation we call *joint direct and transposed sparse matrix-vector multiplication* (SpMM$^{\mathrm{T}}$V). We present a parallel SpMM$^{\mathrm{T}}$V algorithm for shared-memory CPUs. The algorithm uses a sparse matrix format that divides the stored matrix into sparse matrix blocks and compresses the row and column indices of the matrix. This sparse matrix format can be also used for SpMV, SpM$^{\mathrm{T}}$V and similar sparse matrix-vector operations. We expand upon existing research by suggesting new variants of the parallel SpMM$^{\mathrm{T}}$V algorithm and by extending the algorithm to efficiently support symmetric matrices. We compare the performance of the presented parallel SpMM$^{\mathrm{T}}$V algorithm with alternative approaches, which use state-of-the-art sparse matrix formats and libraries, using sparse matrices from real-world applications. The performance results indicate that the median performance of our proposed parallel SpMM$^{\mathrm{T}}$V algorithm is up to 45 % higher than of the alternative approaches. The presented results are a part of [I] .

List of publications:

Submitted:

[I] **Claudio Kozický**, Ivan Šimeček; *Joint Direct and Transposed Sparse Matrix-Vector Multiplication for Multithreaded CPUs*; Concurrency and Computation: Practice and Experience. ISSN 1532-0634.

Doctoral Research Days at FIT 2020

Advances in Information Retrieval from Video Signal

Petr Pulc

Video is a quite fascinating modality. We, people, learn over our first years to understand the environment around us at an incredible level of detail. Sure, we have “stereoscopic” vision, we can manipulate ourselves around the objects to capture their full visual appearance from all sides, or even control some of the items around us to understand them better.

Yet, even in a 2D 15 fps video, the human brain can reconstruct a 3D scene with relative ease. Detect objects and deduct actions. Or most of the time anyway. However, computers fail in many of these tasks miserably. The automotive industry is one of the luckier ones: they can strap virtually any sensor they want on a car. But once we are left with a 2D video, there is no way back.

Deep convolutional neural networks, you say? You wish! And so began my journey on the extraction of human-comprehensible features from a 2D (or, 3D: with colour) video signal.

List of publications:

References

[1] Oliver Kerul-Kmec, Petr Pulc, and Martin Holeňa. “Semisupervised segmentation of UHD video”. In: ITAT. 2018.

[2] Michal Kopp, Petr Pulc, and Martin Holeňa. “Search for Structure in Audiovisual Recordings of Lectures and Conferences”. In: ITAT. 2015.

[3] Petr Pulc and Martin Holeňa. “Application of Meta-learning Principles in Multimedia Indexing”. In: DATESO. 2016.

[4] Petr Pulc and Martin Holeňa. “Case Study in Approaches to the Classification of Audiovisual Recordings of Lectures and Conferences”. In: ITAT. 2014.

[5] Petr Pulc and Martin Holeňa. *Hierarchical Motion Tracking Using Matching of Sparse Features.* 2018.

[6] Petr Pulc and Martin Holeňa. “Towards Real-time Motion Estimation in High-Definition Video Based on Points of Interest”. In: FedCSIS. 2017.

[7] Petr Pulc, Eric Rosenzveig, and Martin Holeňa. “Image Processing in Collaborative Open Narrative Systems”. In: ITAT. 2016.

[8] Petr Pulc et al. *Motion Segmentation by Semi-Supervised Classification in Dynamic Scenery (poster).* 2018.

[9] Tomáš Šabata, Petr Pulc, and Martin Holeňa. “Semi-supervised and Active Learning in Video Scene Classification from Statistical Features”. In: IAL. 2018.

Doctoral Research Days at FIT 2020

User interface for virtual reality

Tomáš Nováček

Virtual reality has been with us for several decades already, but we are still trying to find the right ways to control it. There are a lot of controllers with various purposes and means of input, each with their advantages and disadvantages, but also with specific ways to be handled. Our hands were the primary mean of input for human-computer interaction for a long time. However, now we are able to use movements of our eyes, our feet or event our whole body to control the virtual environment, to interact with it or to move from one place to another. We can achieve that with various controllers and wearable interfaces, like eye tracking, haptic suits or treadmills. One of the biggest problems is still finding a way how to interact with the virtual environment without the need of holding or wearing any device. Our study focuses on controllers based on the motion of hands and fingers of the user, especially with the use of the Leap Motion optical sensor in combination with touchscreens. We created MultiLeap library that combines data from several Leap Motion sensors and provides the most accurate information about the hand's position and pose. To test it, we created a Unity application where data from Leap motions, touchscreens and positional tracking are used to provide even more precise user interaction with the scene.

List of publications:

1. **Tomáš Nováček**, Marcel Jiřina; *Overview of controllers of user interface for virtual reality*; Journal on Multimodal User Interfaces (pending).

Doctoral Research Days at FIT 2020

Sequential Bayesian Modeling of Discrete Random Variables

Radomír Žemlička

Models of discrete counts are popular in many application fields, ranging from the epidemiological data, the number of stock market transactions in finance, traffic intensities in networks and transportation, the number of particle arrivals in physics, to phenomena in social networks [1]. High counts can be generally approximated by continuous data models, but those can fail if the counts are small and include many zeros [4]. Most existing discrete models focus only on static (offline) modeling and disregard the challenging phenomenon of *streaming data*. While the static approaches may be useful if the frequency of data collection is relatively slow, they are doomed to fail in higher frequency scenarios due to their numerical optimization-based nature. The goal of our work is to fill this gap. We aim at computationally cheap methods for online modeling of discrete counts data. These methods should be adaptive, capable to reflect time-varying nature of model parameters, and robust to slight model misspecifications.

Initially, we focused on the Poisson regression model, where the modeled random variable and its linear predictor are linked by the logarithm function. The sequential estimation of the model is, however, generally impossible due to its functional form. It may seem that a repetitively performed static estimation from a data window of a fixed or variable length could solve the problem, but this is prohibitive in many real-time scenarios, mainly due to the intrinsic optimization character [5]. However, the Bayesian paradigm extended by a couple of computationally cheap approximations has been shown to provide a way towards the solution. We utilized El-Sayyad's static Bayesian approach and recasted it to a low-cost sequential algorithm capable of online estimation of regression coefficients [7]. Additionally, knowing that having constant model parameters is rather an exception than a rule, we used the exponential forgetting as a way to deal with potentially slowly time-varying coefficients. Finally, we also wanted to demonstrate the use of the proposed algorithm in the signal processing domain. For this purpose, we opted for the rapidly evolving domain of the distributed inference of unknown variables in networks of cooperating agents. It finds applications in sensor networks, smart grids and microgrids, IoT, big data, social networks, and other types of networked systems [2]. Here, three basic communication and information processing strategies can be distinguished: the incremental strategy, consensus, and diffusion [6]. The point of interest of our research was the diffusion strategy, where the information exchange runs on a single time scale and within one network hop distance without any intermediate iterations [6]. To summarize, we devised a novel framework for sequential (online) distributed inference of the Poisson regression model in networks of collaborating agents. The framework is very scalable. Besides the fully local estimation of possibly time-varying model parameters it allows for efficient collaboration among network agents in order to further accelerate and stabilize the inference [I] .

Currently, our research is focused on the zero-inflated Poisson model, i.e., a model with an excessive number of zeros that cannot be explained by the pure Poisson model. We consider a probabilistic two-component mixture of the Poisson and Dirac distributions. Besides the Poisson model parameters, i.e., the regression coefficients, the new model involves the component probabilities. For the joint estimation of these unknowns we exploit the quasi-Bayesian procedure [3]. This requires knowledge of the posterior predictive distribution of the Poisson distribution, however, it is not analytically tractable. So far, we have had favorable results using the approximation by the negative binomial distribution (NB2), or using the plug-in principle, where the posterior point estimates are used in the generative model.

List of publications:

[I] K. Dedecius and **R. Žemlička**, ”Sequential Poisson Regression in Diffusion Networks,” in *IEEE Signal Processing Letters*, vol. 27, pp. 625-629, 2020, doi: 10.1109/LSP.2020.2987723.

References

[1] N. Bosowski, V. Ingle, and D. Manolakis. "Generalized Linear Models for count time series". In: *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2017, pp. 4272–4276.

[2] Mehmet H Cintuglu and Dmitry Ishchenko. "Secure Distributed State Estimation for Networked Microgrids". In: *IEEE Internet of Things Journal* 6.5 (2019), pp. 8046–8055.

[3] Miroslav Kárný. *Optimized Bayesian Dynamic Advising: Theory and Algorithms (Advanced Information and Knowledge Processing)*. Springer-Verlag London, 2006.

[4] Dimitris Manolakis, Nicholas Bosowski, and Vinay K Ingle. "Count Time-Series Analysis: A Signal Processing Perspective". In: *IEEE Signal Processing Magazine* 36.3 (2019), pp. 64–81.

[5] Raymond H Myers et al. *Generalized linear models (Wiley Series in Probability and Statistics)*. John Wiley & Sons, Mar. 2010. ISBN: 9780470454633.

[6] Ali H Sayed. "Diffusion adaptation over networks". In: *Academic Press Library in Signal Processing*. Vol. 3. Elsevier, 2014, pp. 323–453.

[7] GM El-Sayyad. "Bayesian and classical analysis of Poisson regression". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 35.3 (1973), pp. 445–451.

Doctoral Research Days at FIT 2020

Deep Generative Models and the explainability of neural networks

Jakub Žitný

Deep Generative Models, mainly Generative Adversarial Networks, are often utilized for data synthesis in domains with unbalanced or small datasets. One of these domains is medical imaging — a field where the size and privacy of data as well as model explainability are in demand. We have developed a framework for benchmarking generative models and comparing their performance on various datasets from healthcare. Synthetic data helps improve classification and segmentation models; however, authors rarely use proper evaluation metrics to ensure the quality of synthesized data. We have created a novel formula for such evaluation — one inspired by existing similarity and signal-to-noise metrics, enhanced with heuristics from traditional segmentation tasks. Our new metric correctly identifies synthetic data that should not be used for further model boosting. We believe this metric and the comparison framework overall will lead to better model quality and explainability, especially in medical imaging tasks that are yet to be applied in real-world healthcare.

List of publications:

Accepted:

[I] Jakub Žitný, Pavel Kordík; Evaluation of generative models used for data augmentation in medical imaging domain, ISBI 2020 abstracts

Rejects:

[II] Jakub Žitný, Pavel Kordík; Evaluation of generative models used for data augmentation in medical imaging domain, AIME, 2020

[III] Jakub Žitný, Pavel Kordík; Evaluation of generative models used for data augmentation inmedical imaging domain, AMIA 2020

[IV] Jakub Žitný, Pavel Kordík; An environment for benchmarking generativ emodels in medical imaging, WCMI 2020

Doctoral Research Days at FIT 2020

Dummy Rounds Method as Countermeasure against Side Channel Attacks

Stanislav Jeřábek

The Dummy Rounds protection scheme is intended to offer resistance against Side Channel Attacks (SCA) such as Differential Power Analysis (DPA) [3] to Feistel [2] and Substitution-Permutation [7] ciphers. Its principle is inspired by several well-known countermeasures used in hardware as Hiding [4] and Dynamic Logic Reconfiguration [5] as well as countermeasures used in software implementations as Dummy Cycles [1] or Random Order Execution [8]. The Dummy Rounds method, as we propose, combines software hiding in time with common hardware hiding of the circuitry power consumption. There are more parts of hardware design which are executed, but their outputs are randomly used or not used for computation in every single clock cycle. So, the structure of the design is the same for every clock cycle and also the power consumption stays the same. The final result stays correct due to round scheduling. Experimental evaluation of Dummy Rounds proposed above revealed weaknesses, most notably in the first and last round. The situation can be greatly improved by controlling the transition probabilities in the state space of the algorithm. We derived necessary and sufficient conditions for the round execution probabilities to be uniform and hence the minimum possible. The optimum trajectories over the state space are regular and easy to implement. For the dummy rounds scheme, there is always an optimum set of transition probabilities which makes the round execution probabilities uniform for a particular round now. This ensures maximum resistance against an SCA targeted to a particular round. A trajectory in the optimum set executes a random number of redundant rounds first, then all the active rounds, and then redundant rounds again. Now we are going to use some faster evaluation method than time consuming experimental T-Test [6] evaluation. With the new method, we can quickly evaluate interaction between Dummy Rounds and other known countermeasures.

List of publications:

1. **Stanislav Jeřábek**, Jan Schmidt, Martin Novotný; *Dynamic Reconfiguration as Countermeasure against DPA* In: Proceedings of the Work in Progress Session SEAA/DSD 2017. Linz: Johannes Kepler University, 2017. ISBN 978-3-902457-48-6.

2. **Stanislav Jeřábek**, Jan Schmidt, Martin Novotný, Vojtěch Miškovský; *Dummy Rounds as a DPA countermeasure in hardware* In: Proceedings of the 21st Euromicro Conference on Digital System Design. Piscataway: IEEE, 2018. p. 523-528. ISBN 978-1-5386-7376-8.

3. **Stanislav Jeřábek**, Jan Schmidt; *Analyzing and Optimizing the Dummy Rounds Scheme* In: Proceedings of the 22nd International Symposium on Design and Diagnostics of Electronic Circuits and Systems. Piscataway, NJ: IEEE, 2019. ISBN 978-1-7281-0072-2.

4. Petr Socha, Jan Brejník, **Stanislav Jeřábek**, Martin Novotný, Nele Mentens; *Dynamic Logic Reconfiguration Based Side-Channel Protection of AES and Serpent* In: Proceedings of the 22nd Euromicro Conference on Digital Systems Design. Los Alamitos, CA: IEEE Computer Soc., 2019. p. 277-282. ISBN 978-1-7281-2862-7.

5. Petr Moucha, **Stanislav Jeřábek**, Martin Novotný; *Novel Dummy Rounds Schemes as a DPA Countermeasure in PRESENT Cipher* In: Proceedings of the 23rd International Symposium on Design and Diagnostics of Electronic Circuits and Systems. Piscataway, NJ: IEEE, 2020. p. 1-4. ISBN 978-1-7281-9938-2.

6. Petr Moucha, **Stanislav Jeřábek**, Martin Novotný; *Novel Controller for Dummy Rounds Scheme DPA Countermeasure* In: Proceedings of the 23rd Euromicro Conference on Digital Systems Design. Los Alamitos, CA: IEEE Computer Soc., 2020. p. 281-284. ISBN 978-1-7281-9535-3.

References

[1] Christophe Clavier, Jean-Sébastien Coron, and Nora Dabbous. “Differential Power Analysis in the Presence of Hardware Countermeasures”. In: *Cryptographic Hardware and Embedded Systems — CHES 2000*. Ed. by Çetin K. Koç and Christof Paar. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 252–263. ISBN: 978-3-540-44499-2.

[2] Horst Feistel. “Cryptography and computer privacy”. In: *Scientific american* 228.5 (1973), pp. 15–23.

[3] Paul Kocher, Joshua Jaffe, and Benjamin Jun. “Differential Power Analysis”. In: *Advances in Cryptology — CRYPTO’ 99*. Ed. by Michael Wiener. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 388–397. ISBN: 978-3-540-48405-9.

[4] S. Mangard, E. Oswald, and T. Popp. *Power Analysis Attacks*. 2007, p. 272.

[5] Pascal Sasdrich et al. “Achieving Side-Channel Protection with Dynamic Logic Reconfiguration on Modern FPGAs”. In: *Journal of Cryptographic Engineering* 2 (June 2014), pp. 107–121. ISSN: 2190-8508. DOI: 10.1007/s13389-013-0067-1.

[6] Tobias Schneider and Amir Moradi. “Leakage assessment methodology”. In: *Journal of Cryptographic Engineering* 6.2 (June 2016), pp. 85–99. ISSN: 2190-8516. DOI: 10.1007/s13389-016-0120-y. URL: https://doi.org/10.1007/s13389-016-0120-y.

[7] C. E. Shannon. “Communication theory of secrecy systems”. In: *The Bell System Technical Journal* 28.4 (Oct. 1949), pp. 656–715. ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1949.tb00928.x.

[8] Stefan Tillich, Christoph Herbst, and Stefan Mangard. “Protecting AES Software Implementations on 32-Bit Processors Against Power Analysis”. In: *Applied Cryptography and Network Security*. Ed. by Jonathan Katz and Moti Yung. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 141–157. ISBN: 978-3-540-72738-5.

Doctoral Research Days at FIT 2020

Calculating of Non-Homogeneous Continuous Time Markov Chains

Jan Řezníček

Every system has a lot of components and every component has intensity of failure. Almost every system we could model by the Continuous Time Markov Chains. In my doctoral research I created methods to computing Continuous Time Markov Chains, that has Non-Homogeneous parameters of failures. For this type of computing I used the Matrix Multiplication method by transfering density matrix Q into probability matrix P and than distribution the time *time* into small parts. The next method is inspired by solving the differential equations. Both methods are used to compute the failure distribution function of the system.

List of publications:

1. **Jan Řezníček**, Martin Kohlík, Hana Kubátová; *Hierarchical Dependability Models based on Non-Homogeneous Continuous Time Markov Chains* In: 2019 14th International Conference on Design & Technology of Integrated Systems In Nanoscale Era (DTIS). IEEE, 2019. ISBN 978-1-7281-3424-6.

2. **Jan Řezníček**, Martin Kohlík, Hana Kubátová; *Accurate Inexact Calculations of Non-Homogeneous Markov Chains* In: Proceedings of the 22nd Euromicro Conference on Digital Systems Design. Los Alamitos, CA: IEEE Computer Soc., 2019. p. 470-477. ISBN 978-1-7281-2861-0.

3. **Jan Řezníček**, Martin Kohlík, Hana Kubátová; *Non-homogeneous hierarchical Continuous Time Markov Chains* In: Microprocessors and Microsystems. 2020, 2020(78), ISSN 0141-9331.

4. **Jan Řezníček**, Martin Kohlík, Hana Kubátová; *Non-Homogeneous Continuous Time Markov Chains Calculations* In: Proceedings of the 23rd Euromicro Conference on Digital Systems Design. Los Alamitos, CA: IEEE Computer Soc., 2020. p. 664-671. ISBN 978-1-7281-9535-3.

Doctoral Research Days at FIT 2020

Side-Channel Analysis and Effective Implementations of Embedded Cryptography

Petr Socha

Cryptography has been evolving for a hundreds of years now, as a way to secure confident information against third party. Nowaday cryptographic systems include many diverse embedded devices, such as smartcards used, e.g., for identification or for prepaid services, various IoT applications or even smart cars. While many ciphers currently in use (such as AES) are considered mathematically secure, their implementations may be vulnerable to side channel attacks, such as Differential Power Analysis, Correlation Power Analysis, Mutual Information Analysis and more. These attacks exploit data-dependent power consumption or electromagnetic radiation in order to extract secret/sensitive information from the cryptographic device.

We have examined different approaches to the statistical computations necessary for the first-order [I] and arbitrary-order [IV] side-channel analysis in order to perform these in a numerically stable, parallel and robust fashion, and we have evaluated the memory and time performance of these approaches, as well as their properties regarding practical usage. In order to sucessfully attack implementations in a noisy environment [II], we have proposed and evaluated a novel method for attack evaluation based on a correlation trace derivative [III], [V], which significantly reduces the number of measurements required to mount an attack and in some cases makes the attack even feasible. FPGA-specific polymorphic side-channel countermeasures for AES and Serpent, based on dynamic logic reconfiguration, were proposed and evaluated [VI], [VIII], regarding side-channel leakage, area, and timing, in cooperation with KU Leuven. Furthermore, a high-level synthesis approach to these countermeasures was examined [VII], [XI]. A quite novel concept of distance/energy statistics, in side-channel analysis context, is evaluated and discussed in [X]. Most recently, I have worked with my graduate students on side-channel analysis, effective implementations, and countermeasures, of post-quantum cryptographic algorithms including multivariate quadratic Rainbow signature [IX]. Future work include hardware acceleration of Rainbow on SoC FPGA, its side-channel analysis, countermeasures implementation, and extension of the findings on other post-quantum candidates. I am also a co-author and lecturer of the state-of-the-art side-channel analysis graduate course NI-HSC at CTU in Prague.

List of publications:

Published:

[I] **Socha, P.**; Miškovský, V.; Kubátová, H.; Novotný, M. *Optimization of Pearson correlation coefficient calculation for DPA and comparison of different approaches.* In: Proceedings of the 2017 IEEE 20th International Symposium on Design and Diagnotics of Electronic Circuit & Systems. Piscataway, NJ: IEEE, 2017. p. 184-189. ISSN 2473-2117. ISBN 978-1-5386-0472-4.

[II] **Socha, P.**; Brejník, J.; Bartík, M. *Attacking AES Implementations Using Correlation Power Analysis on ZYBO Zynq-7000 SoC Board.* In: 2018 7th Mediterranean Conference on Embedded Computing (MECO). Piscataway: IEEE, 2018. p. 29-32. ISBN 978-1-5386-5683-9.

[III] **Socha, P.**; Miškovský, V.; Kubátová, H.; Novotný, M. *Correlation Power Analysis Distinguisher Based on the Correlation Trace Derivative.* In: Proceedings of the 21st Euromicro Conference on Digital System Design. Piscataway: IEEE, 2018. p. 565-568. ISBN 978-1-5386-7376-8.

[IV] **Socha, P.**; Miškovský, V.; Novotný, M. *First-Order and Higher-Order Power Analysis: Computational Approaches and Aspects.* In: Proceedings of the 8th Mediterranean Conference on Embedded Computing - MECO'2019. Institute of Electrical and Electronics Engineers, Inc., 2019. p. 83-87. ISSN 2377-5475. ISBN 978-1-7281-1739-3.

[V] **Socha, P.**; Miškovský, V.; Kubátová, H.; Novotný, M. *Efficient algorithmic evaluation of correlation power analysis: Key distinguisher based on the correlation trace derivative.* Microprocessors and Microsystems. 2019, 2019(71), 1-8. ISSN 0141-9331.

[VI] **Socha, P.**; Brejník, J.; Jeřábek, S.; Novotný, M.; Mentens, N. *Dynamic Logic Reconfiguration Based Side-Channel Protection of AES and Serpent.* In: Proceedings of the 22nd Euromicro Conference on Digital System Design. Piscataway: IEEE, 2019. p. 277-282. ISBN 978-1-7281-2862-7.

[VII] **Socha, P.**; Novotný, M. *Towards High-Level Synthesis of Polymorphic Side-Channel Countermeasures.* In: Proceedings of the 23rd Euromicro Conference on Digital System Design. Piscataway: IEEE, 2020. p. 193-199. ISBN 978-1-7281-9535-3.

[VIII] **Socha, P.**; Brejník, J.; Balasch, J.; Novotný, M.; Mentens, N. *Side-channel countermeasures utilizing dynamic logic reconfiguration: Protecting AES/Rijndael and Serpent encryption in hardware.* Microprocessors and Microsystems. 2020, 2020(78), 1-10. ISSN 0141-9331.

Submitted:

[IX] Pokorný, D.; **Socha, P.**; Novotný, M. *Side-channel attack on Rainbow post-quantum signature.* Submitted at Design, Automation and Test in Europe Conference 2021 (DATE'21)

[X] **Socha, P.**; Novotný, M. *A Fair Experimental Evaluation of Distance Correlation Side-Channel Distinguisher.* Submitted at International Symposium on Quality Electronic Design 2021 (ISQED'21)

[XI] **Socha, P.**; Novotný, M. *High-Level Synthesis of Polymorphic Side-Channel Countermeasures for FPGA.* Submitted at Microprocessors and Microsystems journal (MICPRO)

Doctoral Research Days at FIT 2020

SAT Modulo Differential Equation Simulations

Tomáš Kolárik

Formal verification of complex systems, such as embedded systems or cyber-physical systems, is a convenient method to guarantee fulfillment of given specifications. Such systems can often be described by differential equations, in combination with discrete modeling formalisms, such as SAT. Differential equations express physical phenomena of the real world natively and are handled by simulation tools in industry. Such tools, like Simulink, or SpaceEx [1], analyze the systems using exhaustive testing techniques. These approaches, however, lack robust computational support of automatic analysis (e.g., verifying) of such models, and do not search the Boolean state space efficiently, like SAT solvers do. Current solvers, that are based on SAT solvers and that are able to handle differential equations, such as [2], aim at replacing testing techniques by correctness proofs, but use classical mathematical semantics of differential equations. There are fundamental limitations of such state-of-the-art solvers, which inhibit their scalability to practical problems commonly handled by industrial tools. Also, in many applications, classical mathematical semantics often does not correspond well to the actual intended semantics [4]. Thus, we aim on embedding simulation semantics into the SAT framework.

Simulation semantics are based on numerical approaches of solving differential equations [5], which can be very efficient, but cannot guarantee precise bounds of reached error from the exact mathematical solution. This implies that the resulting decision of our method is not exact wrt. classical mathematical semantics, which is quite uncommon in the formal verification community. However, it is exact wrt. simulation semantics, and it seems that the differences of produced results from precise mathematical solutions can be negligible in the models of real systems, with regard to the industrial experience. In the end, there is still no software capable of handling complex practical problems with differential equations, and it is not certain whether precision is something we can not afford here. In the future, we would like to support these observations by robust experiments and theory.

We have already tested our tool [6] in several interesting experiments. Currently, we are working on efficient solving of a quite complex experiment of railway capacity verification, based on [3], which meets potential industrial needs, and also disposes of rich Boolean state space, making it very difficult for the other mentioned tools. In the same time, we are working on improvements of our algorithm, including tighter integration of SAT solving with differential equations, which we believe will increase the efficiency significantly. Solving this experiment will indicate that our method is indeed promising to be used in practical problems.

List of publications:

[I] Tomáš Kolárik, Stefan Ratschan; *SAT Modulo Differential Equation Simulations.* In: Proceedings of 14th International Conference on Tests and Proofs, TAP 2020, held as part of the Software Technologies: Applications and Foundations, STAF 2020, Bergen, Norway, June 22–26, 2020.

References

[1] Goran Frehse et al. “SpaceEx: Scalable Verification of Hybrid Systems”. In: *Computer Aided Verification.* Ed. by Ganesh Gopalakrishnan and Shaz Qadeer. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 379–395. ISBN: 978-3-642-22110-1.

[2] Sicun Gao, Soonho Kong, and Edmund M. Clarke. “dReal: An SMT Solver for Nonlinear Theories over the Reals”. In: *Proceedings of the 24th International Conference on Automated Deduction.* CADE’13. Lake Placid, NY: Springer-Verlag, 2013, pp. 208–214. ISBN: 978-3-642-38573-5. DOI: 10.1007/978-3-642-38574-2_14.

[3] B. Luteberget, K. Claessen, and C. Johansen. “Design-Time Railway Capacity Verification using SAT modulo Discrete Event Simulation”. In: *2018 Formal Methods in Computer Aided Design (FMCAD)*. 2018, pp. 1–9. DOI: 10.23919/FMCAD.2018.8603003.

[4] P. J. Mosterman. “An Overview of Hybrid Simulation Phenomena and Their Support by Simulation Packages”. In: *Lecture Notes in Computer Science*. Vol. 1569. HSCC 1999. Berlin, Heidelberg: Springer-Verlag, 1999. DOI: 10.1007/3-540-48983-5_17.

[5] Kendall Atkinson et al. *Numerical Solution of Ordinary Differential Equations*. John Wiley, Feb. 2009, p. 272. ISBN: 978-0-470-04294-6.

[6] Tomáš Kolárik. *UN/SOT (UN/SAT modulo ODES Not SOT)*. 2020. URL: https://gitlab.com/Tomaqa/unsot.

Index

Řezníček J., 32
Šenkýř D., 15

Beneš T., 17
Bernhauer D., 16
Blažej V., 6
Blizničenko J., 9

Hynek K., 19

Ječmen J., 22
Jeřábek S., 30

Khun J., 20
Knaisl J., 12
Kodýtek F., 23
Kolárik T., 35
Kozický C., 24
Křišťan J.M., 7

Nováček T., 26

Pecka T., 8
Plachý Š., 21

Pulc P., 25

Skotnica M., 10
Slifka J., 11
Socha P., 33
Suchánek M., 13

Červený R., 5

Žemlička R., 27
Žitný J., 29