



**FAKULTA  
INFORMAČNÍCH  
TECHNOLOGIÍ  
ČVUT V PRAZE**

**Výzkumné léto na FIT 2023 (VýLeT 2023):  
Program podpory letního studentského výzkumu na  
FIT ČVUT**

**VYPSANÁ TÉMATA**

## Část 1: Základní informace k přihlášení do programu VýLeT 2023

Tento dokument obsahuje **soupis výzkumných témat vypsaných v rámci programu VýLeT pro rok 2023.**

**Podrobné informace k programu VýLeT 2023 a pravidla programu jsou uvedeny v samostatném dokumentu „Výzkumné léto na FIT 2023 (VýLeT 2023): Program podpory letního studentského výzkumu na FIT ČVUT – PROPOZICE“ zveřejněném na stránce programu na webových stránkách FIT ČVUT.**

### Kdo se může přihlásit?

1. Přihlásit do programu se může student **bakalářského** nebo **navazujícího magisterského** programu na FIT ČVUT, a to v období stanoveném v harmonogramu.
2. Témata se dělí na volná a rezervovaná. Rezervovaná témata jsou určena pro konkrétní studenty vybrané mentorem. Zvolit si rezervované téma může pouze student, pro kterého je téma rezervováno.
3. Kterýkoliv student, který je studentem dle bodu (1) tohoto odstavce se může přihlásit na jakékoliv volné téma.
4. Jeden student se může přihlásit na více témat, vlastní přiřazení zájemců k tématům proběhne až ve výběrovém řízení.
5. Student se může přihlásit a následně řešit výzkumné téma, které je v překryvu s tématem jeho závěrečné práce.
6. Student se může přihlásit a následně řešit výzkumné téma, a posléze si toto téma zvolit jako téma závěrečné práce.
7. Během doby, kdy je program otevřen pro příjem přihlášek od studentů, se student může přihlásit s novým rezervovaným tématem, a to za následujících podmínek: student si najde mentora, dohodne se s ním a následně mentor zadá nové rezervované téma a student se na toto téma přihlásí.
8. Během doby, kdy je program otevřen pro příjem přihlášek od studentů, je možné přidávat témata pouze podle bodu (7) Není možné přidávat témata volná, neurčená pro žádného studenta.

### Jak se přihlásit?

Přihlášení probíhá odesláním elektronického formuláře. Odkaz na formulář je uveden na webu FIT na stránce programu VýLeT 2023. Tento způsob přihlášení je jediný možný a jediný platný. Přihlášení prostřednictvím jiného komunikačního kanálu nebude považováno za platné.

### Do kdy je třeba se přihlásit?

Je třeba se přihlásit ve lhůtě určené pro přihlašování studentů. **Tato lhůta končí 30. 4. 2023.**

## Část 2: Vypsaná výzkumná témata

### Chytré navigační zařízení pro motocykly / Smart navigation devices for motorcycles

**ID zadání / Topic ID: 1**

**Mentor:** prof. Ing. Hana Kubátová, CSc.

**Mentor specialista / Mentor specialist:** Mgr. Jan Balada, ULW Czech, s.r.o.

**Zadání:** V projektu jde o typický problém pro vestavné systémy, kde je důležité navrhovat architekturu hardware i software s ohledem na aplikaci, kde je důležitá velikost, příkon, spotřeba, práce v reálném čase, komunikace i zabezpečení. Navigační zařízení je ve svém základu tablet, na kterém běží operační systém Android. Kompletní funkce ovladačů jsou podporovány aktuálně jen na Android, na iOS běží jen funkce v rámci HID klávesnice. V současné fázi vývoje a používání jde jednak o aktualizaci ROM pro stávající HW na Android 11. Aktuální používaná verze je Android 9. K dispozici je zdrojový kód pro obě verze. Ideální stav je možnost aktualizace systémem OTA, protože Android 9 není pro OTA připraven a oproti Android 11 má jinou strukturu oddílů.

Zároveň začal vývoj kompletně nového hardware, včetně na zakázku vyvíjeného LCD a PCBA. Vývoj PCBA probíhá externě v rámci spolupráce se společností Quectel. Společně s PCBA mají být dodány veškeré součásti operačního systému, kromě aplikační a uživatelské části.

Otevřené problémy do budoucna, které jsou publikovatelné jako velmi inovativní v oblasti automotive aplikací:

1. Externí GNSS prostřednictvím UART a Bluetooth (bude třeba speciální GNSS driver, který zakomponuje především možnost přepnutí na Bluetooth externí GNSS jednotku na systémové úrovni, což je na Androidu obecně problém).
2. Přímé hardwarové připojení na CANBUS;
3. Propojení s periferiemi přes UART;
4. Možnosti bezbateriového provozu.
5. Dálková komunikace (i satelitní), a to jak pro běžné spotřebitele (interkom), tak i rozsáhlejší systémy se zaměřením na prostředí velkých mezinárodních rally.

Očekávané výstupy projektu:

- prezentace výsledku na konferenci o embedded systémech v sekci "praktické aplikace"
- navázání spolupráce s firmou s možností vypracování zajímavé a prakticky uplatnitelné bakalářské nebo diplomové práce
- možná finanční podpora studentů v dalším studiu (PhD) na tématech relevantních k výzkumu KČN

**Assignment:** The project is a typical problem for embedded systems, where it is important to design the hardware and software architecture with regard to the application, where size, power consumption, real-time work, communication and security are important. The navigation device is basically a tablet that runs the Android operating system. The complete functions of the drivers are currently only supported on Android, on iOS only the functions within the HID keyboard run. At the current stage of development and use, it is on the one hand a ROM update for existing HW to Android 11. The current version in use is Android 9. The source code for both versions is available. The ideal situation is to be able to update via OTA, because Android 9 is not OTA-ready and has a different partition structure compared to Android 11.

At the same time, the development of completely new hardware began, including a custom-developed LCD and PCBA. The development of the PCBA takes place externally in cooperation with Quectel. Together with the PCBA, all components of the operating system, except for the application and user parts, are to be delivered.

Open problems for the future that are publishable as very innovative in the field of automotive applications:

1. External GNSS via UART and Bluetooth (a special GNSS driver will be needed, which mainly incorporates the possibility to switch to a Bluetooth external GNSS unit at the system level, which is generally a problem on Android).
2. Direct hardware connection to CANBUS;
3. Connection with peripherals via UART;
4. Possibilities of battery-free operation.
5. Remote communication (including satellite), both for ordinary consumers (intercom) and more extensive systems focusing on the environment of large international rally.

Expected outputs of the project:

- presentation of the result at the conference on embedded systems in the "practical applications" section,
- establishing cooperation with a company with the possibility of developing an interesting and practically applicable bachelor's or diploma thesis,
- possible support in further study (PhD) on topics relevant to KČN research.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- PESW (Prague Embedded System Workshop) [PESW: ČVUT]
- Euromicro Digital System Design (DSD) [Euromicro]

# Vyhodnocení efektivity SAT řešičů pro obvodový SAT / Evaluation of the Efficiency of SAT-Solvers for Circuit-SAT

**ID zadání / Topic ID:** 5

**Mentor:** doc. Ing. Petr Fišer, Ph.D.

**Zadání:** Cílem práce je provést experimentální vyhodnocení efektivity (rychlosti) dostupných open-source řešičů problému splnitelnosti booleovské formule (SAT) pro instance získané transformací z obvodu (netlistu), tj. pro tzv. „circuit-SAT“. Tyto instance jsou svojí povahou specifické. Jsou sice „lehké“ (spíše se blíží 2-SATu), ale objevují se v nich těžké části. SAT řešiče se proto pro ně mohou chovat jinak, než pro běžné zkušební instance. Účinnost dostupných SAT řešičů pro tyto instance zatím nebyla dostatečně zkoumána.

Jedná se o čistě experimentální práci. Detailní znalost problematiky SAT řešičů není nutná. Programování pravděpodobně nebude zapotřebí. Generátor instancí je k dispozici. Naučíte se pracovat s výpočetním clusterem CESNET MetaCentrum (OS Linux).

Výstupem budou příslušné statistiky, doporučení a (pevně doufám) článek na konferenci.

**Assignment:** The aim of this work is to perform an experimental evaluation of the efficiency (speed) of available SAT-solvers for instances obtained by a transformation from a logic circuit (netlist) – the circuit-SAT instances. These instances are somewhat specific. Even thou they are often “easy” to be solved, some difficult parts may appear there. As a result, standard SAT-solvers may behave differently than for typical standard benchmarks. The efficiency of standard SAT-solvers for circuit-SAT hasn't been satisfactorily evaluated yet.

It will be just an experimental work. Detailed knowledge of the SAT-solvers problematic is not needed. Programming probably won't be required either, as the instances generator is available.

You will learn how to work with a national computational cluster CESNET MetaCentrum (OS Linux).

The output of this work will be the respective statistics and recommendations, and hopefully a conference paper.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- IEEE Design and Diagnostics of Electronic Circuits and Systems (DDECS) [IEEE]
- Euromicro Conference on Digital System Design (DSD) [Euromicro]
- The International Conferences on Theory and Applications of Satisfiability Testing (SAT) [TBA]

# Použití logické syntézy pro usnadnění řešení SAT problému / Application of Logic Synthesis to Simplify SAT-Solving

**ID zadání / Topic ID:** 6

**Mentor:** doc. Ing. Petr Fišer, Ph.D.

**Zadání:** Cílem práce je vyzkoušet vliv logické syntézy na řešení problému splnitelnosti booleovské formule (SAT). Tj. prozkoumat, zda a jaká zjednodušení (resp. transformace) booleovské formule vedou ke zrychlení řešení SATu. V literatuře se objevují zmínky o pozitivním i negativním vlivu, dosud ale tento problém nebyl zkoumán dostatečně komplexně.

Jedná se o čistě experimentální práci. Detailní znalost problematiky SAT řešičů a logické syntézy není nutná. Programování pravděpodobně nebude zapotřebí. Naučíte se pracovat s výpočetním clusterem CESNET MetaCentrum (OS Linux) a nástroji pro logickou syntézu.

Výstupem budou příslušné statistiky, doporučení a (pevně doufám) článek na konferenci.

**Assignment:** The aim of this work is to investigate the influence of logic synthesis on SAT-solving. This means, to explore if some simplifications (transformations) of a Boolean formula will lead to a SAT-solving speed-up. Both positive and negative effects were reported in literature, however, this problem hasn't been explored thoroughly enough yet.

It will be just an experimental work. Detailed knowledge of the SAT-solvers problematic and logic synthesis is not needed. Programming probably won't be required either.

You will learn how to work with a national computational cluster CESNET MetaCentrum (OS Linux).

The output of this work will be the respective statistics and recommendations, and hopefully a conference paper.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Design, Automation and Test in Europe (DATE) [IEEE/ACM]
- IEEE Design and Diagnostics of Electronic Circuits and Systems (DDECS) [IEEE]
- The International Conferences on Theory and Applications of Satisfiability Testing (SAT) [TBA]

## Zjišťování podobnosti logických obvodů / Measuring of the Similarity of Logic Circuits

**ID zadání / Topic ID:** 7

**Mentor:** doc. Ing. Petr Fišer, Ph.D.

**Zadání:** Cílem práce je navrhnout algoritmus a naimplementovat nástroj pro stanovení podobnosti dvou kombinačních logických obvodů, které jsou funkčně ekvivalentní, ale strukturně rozdílné. Možností řešení je mnoho, včetně možností definování příslušné metriky. Jedna z možností (možná preferovaná) je navázat na již obhájenou DP, použít zde prezentovaný algoritmus, lépe jej vyhodnotit a případně vylepšit.

Znalost číslicového návrhu není nutná, je to problém spíše z oblasti teoretické informatiky (grafový problém).

Výstupem bude článek na konferenci.

**Assignment:** The aim of this work is to devise an algorithm and implement a tool measuring the similarity of two combinational circuits, which are functionally equivalent, but structurally different. There are many possibilities of doing so, including defining different metrics. Thus, particular solution will be chosen based upon a discussion.

Detailed knowledge of logic synthesis is not required, it is more or less a graph problem.

The output of this work should be a conference paper.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Design, Automation and Test in Europe (DATE) [IEEE/ACM]
- IEEE Design and Diagnostics of Electronic Circuits and Systems (DDECS) [IEEE]
- Euromicro Conference on Digital System Design (DSD) [Euromicro]

## Symbolic Benchmarking / Symbolic Benchmarking

**ID zadání / Topic ID: 8**

**Mentor:** prof. Christoph Kirsch

**Zadání:** The assignment is to study the notion of symbolic benchmarking: given a program  $P$ , possibly as machine code, and a bound  $k$ , find input  $I$  to  $P$  such that  $P$  does not complete execution and/or consume other resources excessively on  $I$  within  $k$  steps, possibly counted as in number of executed machine instructions. The idea is to apply [github.com/cksystemsgroup/unicorn](https://github.com/cksystemsgroup/unicorn) to existing programs and identify performance bugs ideally with program inputs that make the code not complete execution and/or consume other resources such as memory excessively.

Prerequisites: programming languages, algorithms, compilers, runtime systems, architecture, basic mathematical logic

**Assignment:** The assignment is to study the notion of symbolic benchmarking: given a program  $P$ , possibly as machine code, and a bound  $k$ , find input  $I$  to  $P$  such that  $P$  does not complete execution and/or consume other resources excessively on  $I$  within  $k$  steps, possibly counted as in number of executed machine instructions. The idea is to apply [github.com/cksystemsgroup/unicorn](https://github.com/cksystemsgroup/unicorn) to existing programs and identify performance bugs ideally with program inputs that make the code not complete execution and/or consume other resources such as memory excessively.

Prerequisites: programming languages, algorithms, compilers, runtime systems, architecture, basic mathematical logic

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- EuroSys 2024, Fall Deadline: October 2023 [ACM SIGOPS]



## Generování klíčů pomocí fyzicky neklonovatelných funkcí / Generování klíčů pomocí fyzicky neklonovatelných funkcí

**ID zadání / Topic ID: 11**

**Mentor:** Ing. Filip Kodýtek, Ph.D.

**Zadání:** Cílem bude implementace vybrané fyzicky neklonovatelné funkce (PUF) a její následné využití pro generování klíčů. To zahrnuje výběr účelu, pro jaký klíč bude sloužit a následně návrh zpracování výstupu PUF tak, aby jej šlo pro generování klíčů využít.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Design and Diagnostics of Electronic Circuits and Systems (DDECS)
- Digital System Design (DSD)
- Materiály pro výuku předmětů BI-HWB/NI-HWB

# Generátory skutečně náhodných čísel / Generátory skutečně náhodných čísel

**ID zadání / Topic ID: 12**

**Mentor:** Ing. Filip Kodýtek, Ph.D.

**Zadání:** Implementace již existující (případně návrh nové) architektury skutečně náhodného generátoru (TRNG). Analýza vlastností implementovaného TRNG a vyhodnocení jeho náhodnosti.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Digital System Design (DSD)
- Design and Diagnostics of Electronic Circuits and Systems (DDECS)
- Microprocessors and Microsystems (MICPRO)

## Testování Markovských modelů pomocí Monte Carlo simulace na grafické kartě / Testing of Markov models using Monte Carlo simulation on GPU

**ID zadání / Topic ID:** 13

**Mentor:** Ing. Martin Kohlík, Ph.D.

**Mentor specialista / Mentor specialist:** Ing. Jan Řezníček

**Zadání:** V rámci projektu bude vytvořen algoritmus pro testování Markovských modelů pomocí Monte Carlo simulace, který poběží v paralelním prostředí na grafické kartě. Výsledkem práce budou nasimulovaná data, která budou použita pro publikaci článků na konferencích a/nebo časopise. Tento výzkum navazuje na téma disertační práce mentora specialisty a také na výzkum, kterým se zabývá Laboratoř spolehlivosti.

**Assignment:** An algorithm for testing Markov models using Monte Carlo simulation, which will run in a parallel environment on a GPU, will be created within the project. The result of the work will be simulation data, which will be used for the publication of the article at the conferences and/or journal. This research follows up on the topic of the dissertation of a Mentor specialist and also on the research dealt with by the SafetyLab.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- The International Symposium on Design and Diagnostics of Electronic Circuits and Systems (DDECS) [IEEE]
- Conference Series on Digital System Design (DSD) [Euromicro]
- Microprocessors and Microsystems (MICPRO) [Elsevier]

# Techniky strojového učení pro optimalizaci laserového plazmového urychlovače / Machine Learning Techniques for Laser-Plasma Acceleration Optimization

**ID zadání / Topic ID: 14**

**Mentor:** Mgr. Alexander Kovalenko, Ph.D.

**Zadání:** Application of physics Informed machine learning in order to overcome possible challenges by incorporating physical laws and constraints into the machine learning algorithm.

**Assignment:** Application of physics Informed machine learning in order to overcome possible challenges by incorporating physical laws and constraints into the machine learning algorithm.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Scientific Reports Journal

# Bayesovská optimalizace laserového plazmového urychlovače / Bayesian Optimization of a Laser-Plasma Accelerator

**ID zadání / Topic ID:** 15

**Mentor:** Mgr. Alexander Kovalenko, Ph.D.

**Zadání:** Use bayesian statistics to optimize laser-plasma accelerator. Data will be provided.

**Assignment:** Use bayesian statistics to optimize laser-plasma accelerator. Data will be provided.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Scientific Reports Journal

## Efektivní transformery / Efficient transformers

**ID zadání / Topic ID:** 16

**Mentor:** doc. Ing. Pavel Kordík, Ph.D.

**Zadání:** Experimenty se škálovatelností transformerů a dalších podobných ML modelů vzhledem k velikosti kontextového okna. Triky jak použít metalearning a metaoptimalizaci na vytvoření efektivnější struktury paměti.

**Assignment:** Experiments with scalability of transformers and other similar ML models with respect to the size of the context window. Tricks on how to use metalearning and meta-optimization to create a more efficient memory structures.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- ECML PKDD, NEURIPS

## Embedding semantic annotations in Latex / Embedding semantic annotations in Latex

**ID zadání / Topic ID:** 21

**Mentor:** Ing. Milan Dojčinovski, Ph.D.

**Zadání:** While embedding semantic annotations in HTML or XML based formats is already possible via RDFa, microdata or microformats, embedding semantic annotations in Latex is still not well supported. The goal is to develop an approach for embedding semantic annotations (RDF based) in Latex. Part of the thesis is to develop a tool that will support export of the information from latex to RDF/XML.

**Assignment:** While embedding semantic annotations in HTML or XML based formats is already possible via RDFa, microdata or microformats, embedding semantic annotations in Latex is still not well supported. The goal is to develop an approach for embedding semantic annotations (RDF based) in Latex. Part of the thesis is to develop a tool that will support export of the information from latex to RDF/XML.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- SEMANTiCS [Springer]
- ESWC [Springer]
- ISWC [Springer]

# Útok postranním kanálem na statickou spotřebu AES bloku v MCU / Side-channel analysis of the AES MCU crypto block static power

**ID zadání / Topic ID:** 22

**Mentor:** Ing. Jan Bělohoubek, Ph.D.

**Mentor specialista / Mentor specialist:** Ing. Jiří Buček, Ph.D.

**Zadání:** Přípravné práce zahrnují návrh PCB s vybraným procesorem STM32. Návrh PCB může být realizován v úzké spolupráci s vedoucím: začátečníci a zájemci o seznámení se s návrhem PCB jsou vítáni. Dále student připraví jednoduchý software pro MCU a skripty pro řízení experimentu, realizuje experimenty, provede jejich vyhodnocení a bude pod vedením vedoucího pracovat na publikačních výstupech. Experimentální práce je rozdělena do 4 etap. Student realizuje experimenty v etapě (1), v závislosti na postupu prací také v etapách (2), (3) a (4).

Výsledky etapy (1) mohou být publikovány na konferenci jako student paper, úspěch v etapách (2) - (4) umožní podání článku na prestižnější konferenci nebo do časopisu. Etapy (2) - (4) lze provést v libovolném pořadí. V závislosti na postupu prací budeme zvažovat, které etapy a v jakém pořadí realizujeme.

Experimentální část:

(1) Provedení útoku metodou DPA a/nebo CPA na statickou spotřebu AES bloku a srovná složitost útoku s útokem na dynamickou spotřebu. (2) Provedení totožných útoků na MCU, které prošlo standardním procesem simulovaného stárnutí a srovnání s výsledky (1). (3) Provedení totožných útoků na MCU, které obdrželo definovanou dávku ionizujícího záření a srovnání s výsledky (1) a (2). (4) Provedení totožných útoků na rozpouzdřené MCU vystavené neionizujícímu definovanému osvětlení a srovnání s výsledky (1), (2) a (3).

Jako metriku složitosti útoku předpokládáme počet měření nutných pro získání AES klíče.

**Assignment:** The preparatory works for the topic include the design of the PCB. The PCB design will be realized in close cooperation and with help of the topic supervisor: beginners and those who would like to try a PCB design are welcome. Development of the (simple) software for the STM32 MCU, and scripts for conducting experiments follows. The experimental part of the work is further divided into four phases, while phase (1) is mandatory and further phases will follow depending on the success of phase (1). The work will be concluded by experiment evaluation and hopefully a successful publication.

The results of the experimental phase (1) could be already published as a student paper, while success in phases (2) to (4) allows submitting a paper or an article to a prestigious conference or journal. Phases (2) to (4) could be carried out in an arbitrary order, while the phase ordering will be subject to consideration depending on the progress of the work.

Experimental part:

(1) Mounting a DPA and/or CPA attack on the static power of the AES MCU block and comparing its complexity to the attack on the dynamic power of the same block. (2) Mounting the same attacks on the MCU, aged by a standard accelerated aging process, and comparing the attack complexity to the experiments from phase (1). (3) Mounting the same attacks on the MCU, aged by defined ionizing dose radiation, and comparing the attack complexity to the experiments from phases (1) and (2). (4)



Mounting the same attacks on the decapsulated MCU under defined non-ionizing illumination, and comparing the attack complexity to the experiments from phases (1), (2), and (3).

The number of traces is expected as the attack complexity metric is expected.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Euromicro Conference on Digital System Design (DSD) [Euromicro]
- Cryptographic Hardware and Embedded Systems (CHES) [International Association for Cryptologic Research (IACR)]
- Microelectronics Reliability (MICPRO) [Elsevier Ltd.]

# Cluster Ensemble Techniques for Clustering Malware Families / Cluster Ensemble Techniques for Clustering Malware Families

**ID zadání / Topic ID:** 23

**Mentor:** Mgr. Olha Jurečková

**Mentor specialista / Mentor specialist:** Mgr. Martin Jureček, Ph.D.

**Zadání:** In recent years, there are several initiatives in automatic malware categorization using clustering techniques. Clustering algorithms have been applied to different problems in many different real-world applications. However, different clustering algorithms have their own strengths and weaknesses. Therefore, the combination of different clustering algorithms (cluster ensembles) has emerged as an attempt to overcome the limitations of each clustering technique. The use of cluster ensembles aims to combine multiple partitions generated by different clustering algorithms into a single clustering solution (consensus partition). The main objective of the cluster ensemble selection (CES) is the selection of an appropriate subset of base clusterings(BC) and forms a smaller cluster ensemble that performs better than the set of all of the base clusterings. This paper will make an overview of recent research on clustering ensemble and application it in malware detection.

**Assignment:** In recent years, there are several initiatives in automatic malware categorization using clustering techniques. Clustering algorithms have been applied to different problems in many different real-world applications. However, different clustering algorithms have their own strengths and weaknesses. Therefore, the combination of different clustering algorithms (cluster ensembles) has emerged as an attempt to overcome the limitations of each clustering technique. The use of cluster ensembles aims to combine multiple partitions generated by different clustering algorithms into a single clustering solution (consensus partition). The main objective of the cluster ensemble selection (CES) is the selection of an appropriate subset of base clusterings(BC) and forms a smaller cluster ensemble that performs better than the set of all of the base clusterings. This paper will make an overview of recent research on clustering ensemble and application it in malware detection.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- International Conference on Security and Cryptography (SECRYPT). [Institute for Systems and Technologies of Information, Control and Communication (INSTICC)]
- Journal of Computer Virology and Hacking Techniques [Springer]

## Bezpečné aplikace pro automobilový průmysl / Secure automotive applications

**ID zadání / Topic ID:** 26

**Mentor:** prof. Ing. Hana Kubátová, CSc.

**Mentor specialista / Mentor specialist:** Ing. Marek Jánský, Ing. Jakub Horáček, IMA, s.r.o.

**Zadání:** Research work for the IMA company in the field of security, which will be more specifically focused after agreement with the mentors, the topic is also suitable for several students.

The solution will include:

- research on the use of radars and of other sensors for obstacle detection around the vehicle for automotive sensor systems,
- methods of ensuring car connectivity: secure unified gateway, state of development of suitable chips: STM, Infineon, NXP, ...
- vehicle embedded system cyber security.

It is assumed that the results will be used in the context of solving EU projects:

SRIA 2023

<https://www.kdt-ju.europa.eu/documents?data=refresh>

and

Chip Act EU

<https://www.europarl.europa.eu/news/en/headlines/society/20230210STO74502/chips-act-the-eu-s-plan-to-overcome-semiconductor-shortage>

**Assignment:** Research work for the IMA company in the field of security, which will be more specifically focused after agreement with the mentors, the topic is also suitable for several students.

The solution will include:

- research on the use of radars and of other sensors for obstacle detection around the vehicle for automotive sensor systems,
- methods of ensuring car connectivity: secure unified gateway, state of development of suitable chips: STM, Infineon, NXP, ...
- vehicle embedded system cyber security.

It is assumed that the results will be used in the context of solving EU projects:

SRIA 2023

<https://www.kdt-ju.europa.eu/documents?data=refresh>

and

Chip Act EU

<https://www.europarl.europa.eu/news/en/headlines/society/20230210STO74502/chips-act-the-eu-s-plan-to-overcome-semiconductor-shortage>

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Euromicro Conference on Digital System Design (DSD) [EUROMICRO]
- International Symposium on Design and Diagnostics of Electronic Circuits and Systems (DDECS) [IEEE]
- MICPRO journal [Elsevier]

# Velmi líná kompilace pro multi-agentní hledání cest / Very Lazy Compilation for Multi-agent Path Finding

**ID zadání / Topic ID:** 29

**Mentor:** prof. RNDr. Pavel Surynek, Ph.D.

**Zadání:** Při líné kompilaci multi-agentního hledání cest do jiného formalismu, jako je například výroková splnitelnost, určitá omezení zakódujeme, zatímco jiná ponecháme, aby se je řešič naučil sám, tj. jsme při kódování líní. Co kdybychom byli ještě línější, tj. nezakódovali nic, a nechali řešič, ať se kódování problému naučí úplně sám.

**Assignment:** When lazily compiling a multi-agent search for paths to another formalism, such as propositional satisfiability, we encode certain constraints, while we leave others to be learned by the solver himself, ie we are lazy in coding. What if we were even lazier, ie we didn't encode anything, and let the solver learn to encode the problem itself.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- ICAART 2024 [INSTICC]
- ICTAI 2024 [IEEE]

## Resilientní multi-agetní hledání cest s reálnými roboty / Resilient Multi-agent Path Finding with Real Robots

**ID zadání / Topic ID:** 30

**Mentor:** prof. RNDr. Pavel Surynek, Ph.D.

**Zadání:** Představme si vykonávání plánů pro multi-agentní hledání cest se skutečnými roboty, jako jsou roboti, které máme v laboratoři RoboAgeLab, tj. Ozoboti, ePucky a drony Crazyflie. Při vykonávání plánu se občas něco nepovede. Chtěli bychom se zabývat otázkou, jak na nezdar ve vykonávání plánu reagovat změnou plánu, resp. jak zajistit, aby už původní plán byl na takové změny připraven.

**Assignment:** Consider execution of multi-agent path finding plans with real robots, such as the robots we have in the RoboAgeLab lab, ie Ozobots, ePucks and Crazyflie UAVs. Sometimes things go wrong with the plan. We would like to address the question of how to respond to a failure in the execution of the plan by changing the plan or how to ensure that the original plan is ready for such changes.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- ICAART 2024 [INSTICC]
- ICTAI 2024 [IEEE]

# Experimentální studie parametru twin-cover number reálných sociálních sítí / An Experimental Study of the Twin-Cover Number of Real-World Social Networks

**ID zadání / Topic ID:** 32

**Mentor:** Ing. Šimon Schierreich

**Zadání:** Pro mnoho NP-těžkých problémů můžeme ukázat, že, uvažujeme-li platnost rozumných výpočetně složitostních domněnek, nemůže existovat FPT algoritmus pokud parametrizujeme strukturálním parametrem treewidth. V takových případech pak často saháme po jiných strukturálních parametrech, které sice mohou vést k rychlým algoritmům, ale jejich hodnota je malá pro velmi omezené třídy grafů. Klasickým případem takového parametru je vertex cover number.

Cílem této práce je experimentální studie grafového parametru twin-cover number, který je zobecněním zmíněného vertex cover number. Rozdíl mezi oběma parametry spočívá zejména v tom, že, na rozdíl od druhého zmíněného, twin-cover number nabývá nízkých hodnot i pro mnoho tříd hustých grafů. Úkolem studenta je najít vhodné data sety reálných sociálních sítí a na nich zjistit hodnotu tohoto parametru.

**Assignment:** For many NP-hard problems, we can show that, under reasonable computational complexity assumptions, there can be no FPT algorithm if we parameterize by the structural parameter the treewidth of the underlying graph. In such cases, we often turn to other structural parameters that lead to tractable algorithms but are bound only for very restricted graph families. The classic example of such a parameter is the vertex cover number.

The aim of this work is an experimental study of the graph parameter twin-cover number, which is a generalization of the aforementioned vertex cover number. The main difference between the two parameters is that, unlike the latter, the twin-cover number is bounded even for many families of dense graphs. The student's task is to find suitable datasets of real social networks and derive the value of this parameter on these social networks.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- IPEC 2024 [EATCS]
- WAW 2024
- SEA 2024

# Landmark Detection in Micro Robots using Deep Learning / Landmark Detection in Micro Robots using Deep Learning

**ID zadání / Topic ID:** 33

**Mentor:** Mgr. Alexander Kovalenko, Ph.D.

**Zadání:** The task is to detect main features (landmark point, wave variables, etc.) of the micro robots (micro swimmers) from video. Partially labelled dataset will be provided.

**Assignment:** The task is to detect main features (landmark point, wave variables, etc.) of the micro robots (micro swimmers) from video. Partially labelled dataset will be provided.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- International Conference on Computer Vision (ICCV) [The Computer Vision Foundation]
- European Conference on Computer Vision (ECCV) [European Computer Vision Association]



## Témata z oblasti grafových algoritmů, parametrizované složitosti a algoritmů pro těžké problémy / Various topics in the area of graph algorithms, parameterized complexity, and algorithms for hard problems

**ID zadání / Topic ID:** 34

**Mentor:** RNDr. Ondřej Suchý, Ph.D.

**Zadání:** Cílem je zabývat se některými těžkými grafovými problémy z hlediska parametrizace strukturálními parametry. Tedy buď najít parametrizovaný algoritmus pro danou parametrizaci, nebo ukázat, že takový neexistuje. Mezi problémy, na které bychom se mohli zaměřit patří například nalezení geodetického a obalového čísla grafu. Zde podmnožina vrcholů je geodetická, pokud každý vrchol grafu leží na nějaké nejkratší cestě mezi dvěma vrcholy podmnožiny. Geodetické číslo grafu je velikost nejmenší geodetické množiny v grafu. Podmnožina vrcholů grafu je konvexní, když každá nejkratší cesta mezi vrcholy této podmnožiny vede uvnitř této podmnožiny. Obalové číslo grafu je velikost nejmenší množiny vrcholů takové, že jediná konvexní množina, která ji obsahuje, je množina všech vrcholů. Dalším problémem je Geografie, což je hra, kde se hráči střídají v prodluzování cesty v zadaném grafu. Hráč, který cestu nemůže prodloužit, prohrál. Otázkou zůstává, zda první hráč má strategii, která mu zaručí výhru. Mezi vhodnými parametry lze zahrnout například velikost nejmenšího vrcholového pokrytí vstupního grafu, velikost nejmenší feedback vertex set, stromovou sírku, apod. Konkrétní zadání bude stanoveno se studentem.

**Assignment:** The goal is to approach hard graph problems from the perspective of parameterization by structural parameters. That is, to either find a parameterized algorithm with respect to selected parameter, or to show that there is none (under reasonable assumptions). Among the problems that we could focus on is, e.g., the computation of the geodesic and hull number of a graph. Here a subset of vertices is geodetic if every vertex of the graph lies on some shortest path between two vertices of the subset. The geodetic number of a graph is the minimum size of a geodetic set in it. A subset of vertices is convex, if each shortest path between two of its vertices is contained within the subset. The hull number of a graph is the minimum size of a subset of vertices such that the only convex set that this subset is contained in is the whole vertex set of the graph. Yet another interesting problem is Geography, which is a game, where two players alternate in prolonging a path in the input graph. The player unable to prolong the path loses. The question is, given a graph, whether the first player has a winning strategy. Specific goals will be set with the student.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- International Workshop on Graph-Theoretic Concepts in Computer Science + WG [Pro rok 2024 zatím nebyl určen.]
- International Workshop on Combinatorial Algorithms + IWOCA [Pro rok 2024 zatím nebyl určen.]
- Information Processing Letters + IPL [Elsevier]

## The maximum matching game / The maximum matching game

**ID zadání / Topic ID:** 36

**Mentor:** Foivos Fioravantes, Ph.D.

**Assignment:** Consider the following game, which we call "the maximum matching game": two players, Alice and Bob, play on a common graph  $G$ . First, Alice colours a previously uncoloured edge of  $G$  with the colour red, and then Bob colours a previously uncoloured edge of  $G$  blue. Both players know which edges are coloured and by which player during any round of the game. Once all the edges of  $G$  are coloured, the game ends. The score of each player is the order of a maximum matching of  $G$ , formed only by edges coloured by their respective colour. So, for example, if there are 12 edges coloured red, but any matching including only red edges is of order at most 6, then Alice's score is equal to 6. The winner of the game is the player with the highest score. To the best of our knowledge, this is a novel game, that has not been studied before.

The maximum matching game falls into the general family of scoring games, i.e., combinatorial games in which the players strive to achieve the highest score. See [1,2,4,10,11] for other such games that have been recently studied. It is also worth mentioning that, recently, the papers [7,8,9] have started to build a general theory around scoring games.

Assuming both players of the maximum matching game play in an "optimal" way, what interests us is to answer questions like:

- 1) Could Bob ever win this game?
- 2) Given an arbitrary graph  $G$ , can we efficiently (in polynomial time) decide which player will win?
- 3) Should Bob follow a greedy strategy, each step picking an edges that maximises his own score, or should he try to stop Alice from achieving a high score herself?
- 4) If the answer to question 2. is "no", then can we answer to the same question but when focusing only on specific graph classes?

Typically, deciding the outcome of such games tends to be PSPACE-complete. However, given the fact that the game we propose deals with the edges of a graph, and the goal is to build a matching, we would not be extremely surprised if the answer to question 2. above turned out to be "yes".

It could also be interesting to consider the Maker-Breaker version of the above game: this version is played in the same way, the difference being that now Bob disregards his own score, and has as a goal to minimise Alice's score. In a way, this would model the scenario where we are trying to construct a maximum matching of a graph, and roughly half the edges are chosen in the worse possible way. Maker-Breaker games are, generally, well studied, and form an important subcategory of Positional Games [6]. Note that, judging from our results on the maximum matching game, the Maker-Breaker version could be the subject of a dedicated study.

What we expect:

- 1) A short literature review of games similar to the one we propose. For example, the authors of [3,5] study combinatorial games revolving around matchings, with rather different rules than the game we propose. Nevertheless, these games could share some similar behaviours with the maximum matching game.

2) Answering (at least some of) the questions asked above, as well as any new questions that can arise while studying the game we propose.

## References

- [1] J. Bensmail, F. Fioravantes, F. Mc Inerney, and N. Nisse. The largest connected subgraph game. *Algorithmica*, 84(9):2533–2555, 2022.
- [2] M. Changat, D. S. Lekha, I. Peterin, A. R. Subhamathi, and S. Spacapan. The median game. *Discrete Optimization*, 17:80-88, 2015.
- [3] D. W. Cranston, W. B. Kinnersley, S. O, and D. B. West. Game matching number of graphs. *Discrete Applied Mathematics*, 161(13):1828–1836, 2013.
- [4] E. Duchêne, S. Gonzalez, A. Parreau, E. Rémila, and P. Solal. INFLUENCE: a partizan scoring game on graphs. *Theoretical Computer Science*, 878-879:26-46, 2021.
- [5] W. Goddard and M. A. Henning. The matcher game played in graphs. *Discrete Applied Mathematics*, 237:82–88, 2018.
- [6] D. Hefetz, M. Krivelevich, M. Stojaković, and T. Szabó. *Positional Games*, volume 44 of Oberwolfach Seminars. Springer, Basel, 2014.
- [7] U. Larsson, R. J. Nowakowski, J. P. Neto, and C. P. Santos. Guaranteed scoring games. *Electronic Journal of Combinatorics*, 23(3): P3.27, 2016.
- [8] U. Larsson, R. J. Nowakowski, and C. P. Santos. Game comparison through play. *Theoretical Computer Science*, 725:52-63, 2018.
- [9] U. Larsson, R. J. Nowakowski, and C. P. Santos. Games with guaranteed scores and waiting moves. *International Journal of Game Theory*, 47(2):653-671, 2018.
- [10] P. Micek and B. Walczak. A graph-grabbing game. *Combinatorics, Probability & Computing*, 20(4):623-629, 2011.
- [11] A. Shapovalov. Occupation games on graphs in which the second player takes almost all vertices. *Discrete Applied Mathematics*, 159(15):1526-1527, 2011.

## **Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- 50th International Workshop on Graph-Theoretic Concepts in Computer Science. WG 2024. [University of Primorska, in cooperation with the Slovenian Discrete and Applied Mathematics Society.]
- *Discrete Applied Mathematics*. [Elsevier]
- *Theoretical Computer Science*. [Elsevier]

# Optimalizace uživatelské zkušenosti v architektuře orientované na služby: Shromažďování a využití zpětné vazby od zákazníků / Optimizing User Experience in Service-Oriented Architecture: Gathering and Utilizing Customer Feedback

**ID zadání / Topic ID:** 37

**Mentor:** Ing. Marek Suchánek

**Zadání:** Cílem této studie je optimalizovat uživatelskou zkušenost v prostředí architektury orientované na služby (SOA) prostřednictvím sběru a využití zpětné vazby od zákazníků. V SOA spolupracuje více služeb, aby poskytovaly bezproblémové uživatelské prostředí. Zajištění toho, že všechny služby fungují optimálně a splňují potřeby uživatelů, však může být náročné. Aby se tato studie vypořádala s tímto problémem, zaměří se na získávání zpětné vazby od uživatelů ohledně jejich zkušeností s různými službami a využití této zpětné vazby ke zlepšení. Studie bude zahrnovat vývoj mechanismu zpětné vazby, který uživatelům umožní snadno poskytovat zpětnou vazbu, analýzu zpětné vazby k identifikaci vzorců a oblastí pro zlepšení a spolupráci s vývojovým týmem na implementaci nezbytných změn. Tato studie navíc navrhne návrh univerzálních implementací mechanismu zpětné vazby, který lze aplikovat na různá prostředí SOA. Upřednostňováním zpětné vazby od uživatelů a neustálým zlepšováním si tato studie klade za cíl zlepšit uživatelskou zkušenost a zlepšit spokojenost zákazníků v prostředí architektury orientované na služby. Výsledky této studie přispějí k souboru znalostí o zpětné vazbě uživatelů v SOA, poskytnou praktická doporučení pro zlepšení uživatelské zkušenosti v systémech orientovaných na služby a navrhnou návrh univerzálních implementací mechanismu zpětné vazby.

**Assignment:** The objective of this study is to optimize the user experience in a service-oriented architecture (SOA) environment through the collection and utilization of customer feedback. In SOA, multiple services collaborate to deliver a seamless user experience. However, ensuring that all services are functioning optimally and meeting user needs can be challenging. To address this challenge, this study will focus on gathering feedback from users on their experience with the different services and using that feedback to make improvements. The study will involve developing a feedback mechanism that enables users to provide feedback easily, analyzing the feedback to identify patterns and areas for improvement, and collaborating with the development team to implement necessary changes. Additionally, this study will propose a design for versatile implementations of the feedback mechanism that can be applied to different SOA environments. By prioritizing user feedback and making continuous improvements, this study aims to enhance the user experience and improve customer satisfaction in a service-oriented architecture environment. The results of this study will contribute to the body of knowledge on user feedback in SOA, provide practical recommendations for improving user experience in service-oriented systems, and propose a design for versatile implementations of the feedback mechanism.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Information Systems Journal [Wiley]
- Evaluation of Novel Approaches to Software Engineering [INSTICC]
- International Conference on Software Technologies [INSTICC]

## Paralelní vícecestný Quicksort (dokončení a publikace) / Parallel multi-way Quicksort

**ID zadání / Topic ID:** 38

**Mentor:** doc. Ing. Ivan Šimeček, Ph.D.

**Mentor specialista / Mentor specialist:** Ing. Daniel Langr, Ph.D.

**Zadání:** O tématu "Paralelní vícecestný Quicksort" byla vypracována nadstandardní práce <https://projects.fit.cvut.cz/theses/4396> Je tedy k dispozici jak text tak zdrojové kódy. Práce byla rozhodně zasloužila dovést k publikaci. To by zhruba znamenalo (samozřejmě s mou pomocí): identifikace důležitých pasáží + případně dopsání dalších (minimum), překlad, další generatory vstupních dat, měření vlivu kompilátoru, měření na více platformách, porovnání s více "konkurenty", vyhodnocení apod.

**Assignment:** O tématu "Paralelní vícecestný Quicksort" byla vypracována nadstandardní práce <https://projects.fit.cvut.cz/theses/4396> Je tedy k dispozici jak text tak zdrojové kódy. Práce byla rozhodně zasloužila dovést k publikaci. To by zhruba znamenalo (samozřejmě s mou pomocí): identifikace důležitých pasáží + případně dopsání dalších (minimum), překlad, další generatory vstupních dat, měření vlivu kompilátoru, měření na více platformách, porovnání s více "konkurenty", vyhodnocení apod.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Časopis zabývající se paralelním zpracováním např.: Journal of Parallel and Distributed Computing, International Journal of Parallel, Emergent and Distributed Systems, Scalable Computing: Practice and Experience apod. [uvidíme viz výše]

# Approximation and Parameterization of Max Min Cycle Transversal Problems / Approximation and Parameterization of Max Min Cycle Transversal Problems

**ID zadání / Topic ID:** 39

**Mentor:** Nikolaos Melissinos, Ph.D.

**Assignment:** One of the most classic problems in theoretical computer science is the Feedback Vertex Set Problem (FVS). Given a graph  $G$ , a set of vertices  $S$  is called feedback vertex set if by removing  $S$  from  $G$ , we obtain an acyclic graph.

In FVS we are given a graph  $G$  and asked to find a feedback vertex set of minimum size. FVS belongs to the general class of circular transversal problems. Two other problems in this category are Odd Cycle Transversal (OCT) and Even Cycle Transversal (ECT). Both of these problems are similar to FVS except that in OCT the resulting graph contains no odd cycles while in ECT it contains no even cycles.

A variation of FVS, called Maximum Minimal FVS (MMFVS) asks to find a feedback vertex set  $S$  of the given graph which is also minimal (i.e. no proper subset of  $S$  is a feedback vertex set). In [1] the authors consider the approximability of MMFVS, while in [2] and [3] the problem is been studied under the lens of parameterized complexity.

Here, we propose the study of the approximability and parameterization of Maximum Minimal OCT and Maximum Minimal ECT. First we will consider the approximability of the two problems. Our goal is to develop polynomial-time approximation algorithms and find lower bounds on the approximation ratio that can be achieved in polynomial time. Next, we will study the parameterized complexity of the problems based on parameters such as the treewidth of the given graph and the size of the solution. We will also attempt to prove the optimality of our parameterized algorithms by finding lower bounds on their running time under standard assumptions, such as the exponential time assumption.

[1] Dublois, L., Hanaka, T., Khosravian Ghadikolaei, M., Lampis, M., Melissinos, N.:

(In)approximability of maximum minimal FVS. J. Comput. Syst. Sci. 124, 26–40

(2022). <https://doi.org/10.1016/j.jcss.2021.09.001>

[2] Gaikwad, A., Kumar, H., Maity, S., Saurabh, S., Tripathi, S.K.: Maximum minimal

feedback vertex set: A parameterized perspective. CoRR abs/2208.01953 (2022).

<https://doi.org/10.48550/arXiv.2208.01953>

[3] Lampis, M., Melissinos, N., Vasilakis, M.: Parameterized Max Min Feedback Vertex

Set. CoRR abs/2302.09604 (2023). <https://doi.org/10.48550/arXiv.2302.09604>

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- 50th International Workshop on Graph-Theoretic Concepts in Computer Science - WG 2024 [University of Primorska, Kranjska Gora, Slovenia]
- The 35th International Workshop on Combinatorial Algorithms, IWOCA 2023 [not announced yet]
- Theoretical Computer Science [Elsevier]

## Analýza C++ kódu pro vědecké výpočty / Analysis of C++ code for scientific computations

**ID zadání / Topic ID:** 40

**Mentor:** doc. Ing. Ivan Šimeček, Ph.D.

**Zadání:** Na adrese <https://github.com/vincefn/objcryst> jsou C++ zdrojové kódy programu v oblasti krystalografie. Cílem tohoto výzkum je analýza jedné klíčové fce GetLogLikelihood() a návrh na její možnou paralelizaci tzn. analýza datových struktur, jejich závislostí apod.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Výsledek v ideálním případě bude přímo publikovatelný v krystalografických časopisech jako: Journal of Applied Crystallography apod. [Uvidíme dle výsledku.]



# Differentiable Surface Triangulation for Quad-Dominant Surface Meshes / Differentiable Surface Triangulation for Quad-Dominant Surface Meshes

**ID zadání / Topic ID:** 41

**Mentor:** Mgr. Vojtěch Rybář

**Zadání:** Extend the differentiable surface triangulation method introduced by Rakotosaona et al. to generate quad-dominant meshes for 3D surfaces by developing a new formulation of the mesh generation process or modifying the loss function.

Quad-dominant meshes offer several advantages over triangular meshes in various applications, which motivate the need to extend the differentiable surface triangulation method to handle them.

Some of the key areas where quad-dominant meshes are used include:

- \* Computer Graphics: Quad elements can provide smoother and more continuous surfaces compared to triangles, especially when using subdivision surface algorithms. These algorithms usually subdivide quads into smaller quads, maintaining the smoothness and continuity of the original mesh. This smoothness is essential for creating realistic and visually appealing 3D models.

- \* Finite element analysis (FEA): Quad elements are often preferred in FEA due to their improved numerical stability and convergence properties compared to triangular elements.

By extending the differentiable surface triangulation method to quad-dominant meshes, we can leverage these advantages in a wide range of applications, making the method more versatile and beneficial for end-to-end machine learning frameworks in geometry processing tasks.

Reference:

RAKOTOSAONA, Marie-Julie, et al. Differentiable surface triangulation. ACM Transactions on Graphics (TOG), 2021, 40.6: 1-13.

**Assignment:** Extend the differentiable surface triangulation method introduced by Rakotosaona et al. to generate quad-dominant meshes for 3D surfaces by developing a new formulation of the mesh generation process or modifying the loss function.

Quad-dominant meshes offer several advantages over triangular meshes in various applications, which motivate the need to extend the differentiable surface triangulation method to handle them.

Some of the key areas where quad-dominant meshes are used include:

- \* Computer Graphics: Quad elements can provide smoother and more continuous surfaces compared to triangles, especially when using subdivision surface algorithms. These algorithms usually subdivide quads into smaller quads, maintaining the smoothness and continuity of the original mesh. This smoothness is essential for creating realistic and visually appealing 3D models.

- \* Finite element analysis (FEA): Quad elements are often preferred in FEA due to their improved numerical stability and convergence properties compared to triangular elements.

By extending the differentiable surface triangulation method to quad-dominant meshes, we can leverage these advantages in a wide range of applications, making the method more versatile and beneficial for end-to-end machine learning frameworks in geometry processing tasks.

Reference:

RAKOTOSAONA, Marie-Julie, et al. Differentiable surface triangulation. ACM Transactions on Graphics (TOG), 2021, 40.6: 1-13.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Eurographics [ACM SIGGRAPH]
- Symposium on Geometry Processing [VCI RWTH Aachen University]

# Analýza aproximativních implementací neuronových sítí na FPGA / Analysis of approximative implementations of neural networks na FPGA

**ID zadání / Topic ID:** 42

**Mentor:** Ing. Miroslav Skrbek, Ph.D.

**Zadání:** Předmětem práce je analýza aproximativních implementací neuronových sítí vytvořených vysokoúrovňovou syntézou z jazyka C do VHDL a následnou syntézou do FPGA. Projekt vychází z integrace aproximativního počítání do nástroje HLS4ML, který z Keras modelu generuje vstup pro vysokoúrovňovou syntézu. Cílem práce je vytvoření sady implementací s různými velikostmi a přesnostmi výpočtů. Výsledky pak srovnat z hlediska potřebných zdrojů FPGA, rychlosti výpočtu, spotřeby, případně dalších parametrů, které se dají získat z vývojových nástrojů.

**Assignment:** The subject of the project is the analysis of approximate implementations of neural networks created by high-level synthesis tool from the C language to VHDL and subsequent synthesis to FPGA. The project is based on the integration of approximate computing into the HLS4ML tool, which generates input for high-level synthesis from a Keras model. The aim of the work is to create a set of implementations that differ in size and calculation accuracies. Then to compare them in terms of the necessary FPGA resources, calculation speed, chip consumption, or other parameters that can be obtained from development tools.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- DSD: Digital Systems Design 2024
- MECO Mediterranean Conference on Embedded Computing 2024

## Improved denoising diffusion probabilistic models with attention for super-resolution / Improved denoising diffusion probabilistic models with attention for super-resolution

**ID zadání / Topic ID:** 44

**Mentor:** Mgr. Petr Šimánek

**Zadání:** Super-resolution of weather data is a very important topic. Numerical weather simulations are very time consuming and especially complicated in high resolution. We want to improve the spatial resolution by denoising diffusion probabilistic models. Such models were used recently, but the results are worse than the results of Generative Adversarial Networks. We will improve the current state-of-the-art by:

1. Using stable diffusion: Replace the standard isotropic diffusion process with stable diffusion, which uses an anisotropic noise distribution. This can lead to faster convergence and better sample quality.
2. Using attention mechanisms: Integrate attention layers into the de-noising model to help it focus on the most relevant parts of the input data. This can be done using self-attention layers or Transformer architectures.
3. Modifying loss function: Incorporate domain-specific knowledge or constraints into the loss function to guide the learning process towards more physically plausible weather patterns.

The output will be novel variant of DDPM for super-resolution of weather data.

**Assignment:** Super-resolution of weather data is a very important topic. Numerical weather simulations are very time consuming and especially complicated in high resolution. We want to improve the spatial resolution by denoising diffusion probabilistic models. Such models were used recently, but the results are worse than the results of Generative Adversarial Networks. We will improve the current state-of-the-art by:

1. Using stable diffusion: Replace the standard isotropic diffusion process with stable diffusion, which uses an anisotropic noise distribution. This can lead to faster convergence and better sample quality.
2. Using attention mechanisms: Integrate attention layers into the de-noising model to help it focus on the most relevant parts of the input data. This can be done using self-attention layers or Transformer architectures.
3. Modifying loss function: Incorporate domain-specific knowledge or constraints into the loss function to guide the learning process towards more physically plausible weather patterns.

The output will be novel variant of DDPM for super-resolution of weather data.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- ICLR - International Conference on Learning Representations

## Improved denoising diffusion probabilistic models with attention for super-resolution / Improved denoising diffusion probabilistic models with attention for super-resolution

**ID zadání / Topic ID:** 45

**Mentor:** Mgr. Petr Šimánek

**Zadání:** In this task, the goal is to investigate the behavior of neural cellular automata (NCA), a form of artificial life, by employing evolutionary optimization techniques. The primary objective is to train NCAs to maximize certain metrics related to complex system properties, such as causal emergence, information integration, transfer entropy, and agentness. Causal emergence measures the degree to which macro-level behavior emerges from micro-level interactions, while information integration quantifies the extent to which a system combines and processes information. Transfer entropy evaluates the flow of information between different parts of a system, and agentness assesses the degree to which an entity exhibits goal-directed behavior. By optimizing these metrics, we aim to gain a deeper understanding of the underlying principles governing NCA behavior and harness their potential for various applications in artificial life research. The output of the work is the training of such NCAs and analysis of the results.

**Assignment:** In this task, the goal is to investigate the behavior of neural cellular automata (NCA), a form of artificial life, by employing evolutionary optimization techniques. The primary objective is to train NCAs to maximize certain metrics related to complex system properties, such as causal emergence, information integration, transfer entropy, and agentness. Causal emergence measures the degree to which macro-level behavior emerges from micro-level interactions, while information integration quantifies the extent to which a system combines and processes information. Transfer entropy evaluates the flow of information between different parts of a system, and agentness assesses the degree to which an entity exhibits goal-directed behavior. By optimizing these metrics, we aim to gain a deeper understanding of the underlying principles governing NCA behavior and harness their potential for various applications in artificial life research. The output of the work is the training of such NCAs and analysis of the results.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- GECCO - The Genetic and Evolutionary Computation Conference

# Fuzzy porovnávání klíčových slov a frází pro rozpoznávání pojmenovaných entit / Fuzzy Keyword-matching and Phrase-matching for Named Entity Recognition

**ID zadání / Topic ID:** 48

**Mentor:** Ing. Vojtěch Cahlík

**Zadání:** Porovnávání klíčových slov je jednoduchou, nicméně mocnou technikou rozpoznávání pojmenovaných entit (named entity recognition) v přirozeném jazyce, kterou lze použít buď samostatně, nebo jako podpůrnou metodu pro systémy založené na modelech hlubokého učení. V rámci tomto přístupu je pro každou kategorii pojmenovaných entit (např. "bakterie") k dispozici seznam souvisejících klíčových slov (např. "Salmonella", "Streptococcus"), který se poté využije pro porovnávání slov v dotazovaných textech.

Toto téma se zaměřuje na výzkum a vývoj pokročilých technik porovnávání klíčových slov, zejména porovnávání frází (např. "Escherichia coli") a fuzzy porovnávání (např. "E. coli"). Přístupů k prozkoumání se nabízí celá řada, kupříkladu metody založené na podobnosti (příkladem je editační vzdálenost a embeddingy slov či frází) nebo techniky používané v příbuzné problematice word alignmentu [1]. Tyto přístupy budou v rámci výzkumu kvalitativně porovnány. Důležitým faktorem je výpočetní efektivita, poněvadž počet kandidátních frází při zvyšující se délce textu exploduje.

Výstupem studentova výzkumu bude samostatná vědecká publikace. Získané poznatky budou dále využity výzkumníky v DataLabu na reálném projektu "Analýza textových dat – extrakce entit".

[1] <https://arxiv.org/abs/2004.08728>

**Assignment:** Keyword matching is a simple but powerful named entity recognition technique in natural language text, which can be used by itself, or to support systems based on deep learning models. With this approach, for each named entity category (e.g. "bacteria"), a list of related keywords (e.g. "Salmonella", "Streptococcus") is provided, which is then used for the matching of words in query texts.

This topic focuses on the research and development of advanced keyword-matching techniques, such as phrase matching (e.g. "Escherichia coli") or fuzzy matching (e.g. "E. coli"). Numerous approaches can be investigated, such as similarity-based methods (e.g. edit distance, word embeddings, or phrase embeddings), or techniques used in the related problem of word alignment [1]. These approaches will be qualitatively compared. An important factor is computational efficiency, as the number of candidate phrases explodes with longer texts.

The output of the student's research will be a stand-alone scientific publication. The insights will be further utilized by researchers in DataLab in the real-world project "Textual data analysis - entity extraction".

[1] <https://arxiv.org/abs/2004.08728>

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD) [The steering committee of ECML PKDD]

## Automatické vyhodnocení testů manuální zručnosti podle videozáznamu / Automatic evaluation of manual dexterity tests based on video footage

**ID zadání / Topic ID:** 49

**Mentor:** Ing. Tomáš Vondra, Ph.D.

**Mentor specialista / Mentor specialist:** Ing. Jakub Novák

**Zadání:** Na Klinice rehabilitačního lékařství 1.LF UK a VFN v Praze probíhá výzkum zaměřený na standardizované testy hodnotící funkcí horních končetin. Jedním z nich je Box and Block Test, kterým se hodnotí obratnost rukou pacientů. Test spočívá v co nejrychlejším přemístování kostek z krabice přes přepážku do druhé krabice tak, aby se pacient konečky prstů držících přemísťovanou kostku dostal za přepážku, než kostku upustí. Výsledkem je počet kostek jednotlivě přemístěných v časovém limitu za dodržení instrukcí. Vyhodnocování se v současné době provádí v reálném čase s pacientem, s kontrolou druhým hodnotitelem podle videa. Vyhodnocení je velmi náročné na pozornost hodnotitelů a náchylné na chyby. Navrhněte způsob, jak automaticky vyhodnocovat tento test pomocí technik strojového vidění. Srovnajte výsledky s lidskými hodnotiteli. Práce je interdisciplinární; výsledek je možné uplatnit jak na konferencích na téma umělé inteligence, tak těch rehabilitačních. Výsledný program by se mohl využívat v klinické praxi rehabilitačních zařízení, která tento test běžně používají, a odstranil by potřebu druhého hodnotitele.

**Assignment:** The Department of Rehabilitation Medicine of the First Faculty of Medicine, Charles University and the General Hospital in Prague is conducting research focused on standardized tests evaluating the function of the upper limbs. One of them is the Box and Block Test, which assesses the dexterity of patients' hands. The test consists of moving the dice from the box across the partition to the second box as quickly as possible so that the patient must get over the partition with the fingertips holding the moved dice before dropping the dice. The result is the number of dice individually moved within the time limit while following the instructions. The evaluation is currently performed in real time with the patient, with the reevaluation of a second evaluator using video. The evaluation is very demanding on the attention of evaluators and prone to errors. Suggest a way to automatically evaluate this test using machine vision techniques. Compare the results with human evaluators. The work is interdisciplinary; the result can be applied both at conferences on the topic of artificial intelligence and rehabilitation ones. The resulting program could be used in the clinical practice of rehabilitation facilities that commonly use this test, and would remove the need of the second evaluator.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Database ECML PKDD [Nezávislá konference / Springer]
- Rehabilitace a fyzikální lékařství [Česká lékařská společnost J. E. Purkyně]

## Modelování výkonu nad algoritmy strojového učení / Performance modeling over machine learning algorithms

**ID zadání / Topic ID:** 51

**Mentor:** Ing. Tomáš Vondra, Ph.D.

**Zadání:** Máme již výsledky v oblasti predikce zátěže typu webových služeb a chtěli bychom je aplikovat na algoritmy strojového učení, jako jsou NLU (Natural Language Understanding). Tyto pracují ve fázi učení na velkých souborech dat a jejich doba běhu je úměrná velikosti dat. Lze jí zmenšit přidáním více výpočetních uzlů, kde zrychlení může, ale nemusí být lineární v závislosti na algoritmu a distribuci vstupních dat do uzlů. Bylo by zajímavé vytvořit modely tohoto vztahu a umět předpovídat dobu běhu na základě velikosti dat a velikosti clusteru. Umožnilo by to vědci v roli cloudového zákazníka vytvořit klastr dostatečně velký na to, aby spočítal jeho práci, například za jednu hodinu.

Ve fázi predikce se pak model chová jako služba, která vrácí předpovědi, a má klasické problémy jako je rychlost odezvy, paralelizace a dynamické škálování.

V úloze budou použita data z reálného provozu v ČSOB.

**Assignment:** We already have results in web service type load prediction and would like to apply them to machine learning algorithms such as NLU (Natural Language Understanding). These work in the learning phase on large data sets and their running time is proportional to the size of the data. It can be reduced by adding more compute nodes, where the speedup may or may not be linear depending on the algorithm and the distribution of input data to the nodes. It would be interesting to model this relationship and be able to predict runtime based on data size and cluster size. It would allow a scientist in the role of a cloud customer to create a cluster large enough to count his work in, say, one hour.

In the prediction phase, the model behaves like a service that returns predictions and has classic problems such as response speed, parallelization and dynamic scaling.

The task will use data from real operation at ČSOB.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- International Conference on Advanced Computer Information Technologies (ACIT) [West Ukrainian National University; indexed by IEEE]



## Rozpoznání textu v archaických písmech / Recognition of text in archaic scripts

**ID zadání / Topic ID:** 52

**Mentor:** Ing. Tomáš Vondra, Ph.D.

**Zadání:** Zadavatelem je CATVUSA, nezisková televize propagující ČR v USA. Cílem práce je rozšířit WP plugin Genealogy na našem webu tím, že by tam byla další funkce, která by dokázala přepsat německý text tištěný ve fontu Kurent a Schwabach do latinky a pokud to bude možné i přeložit do angličtiny za použití například Google translate. Prostě že by někdo oscanoval text v kurentu (třeba staré noviny) a pak by přes náš plugin získal tento text v latině a popřípadě přeložený do angličtiny.

Tento plugin chceme rozšířit o výše uvedenou funkci se čtením scanu a s překladem <https://catvusa.com/genealogy/german-czech/>

Předpokládáme, že by se zadání dalo řešit zabudováním open-source OCR enginu a jeho naučením na archaické znakové sady.

**Assignment:** The original submitter is CATVUSA, a non-profit television promoting the Czech Republic in the USA. The goal of the work is to expand the Genealogy WP plugin on our website by adding another function that would be able to transcribe the German text printed in the Kurent and Schwabach font into Latin and, if possible, also translate it into English using, for example, Google translate. It's just that someone would scan the text in the kurent (for example, an old newspaper) and then via our plugin, they would get this text in Latin and possibly translated into English.

We want to expand this plugin with the above function with scan reading and translation <https://catvusa.com/genealogy/german-czech/>

We assume that the assignment could be solved by building in an open-source OCR engine and teaching it to archaic character sets.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- Conference Information Technologies – Applications and Theory (ITAT) [P.J.Šafárik University in Košice, Slovakia]

# Multi-Agent Reinforcement Learning with MAPPO and Social Influence Reward / Multi-Agent Reinforcement Learning with MAPPO and Social Influence Reward

**ID zadání / Topic ID:** 53

**Mentor:** Mgr. Vojtěch Rybář

**Zadání:** Multi-Agent Proximal Policy Optimization (MAPPO) is an advanced algorithm for multi-agent reinforcement learning. Social Influence Reward (SIR) is an innovative concept proposed by Jacques et al. (2019) that enhances the learning dynamics and adaptability of agents in multi-agent systems. The success of MAPPO in cooperative games, as demonstrated by Vinitzky et al. (2022), offers valuable insights for improving the performance of MAPPO.

Incorporating SIR into MAPPO can enhance agent cooperation and adaptability, which are crucial aspects of multi-agent systems. The success of MAPPO in cooperative games (Vinitzky et al., 2022) suggests that integrating SIR into MAPPO could lead to interesting improvements in multi-agent reinforcement learning.

The goal is to incorporate SIR into MAPPO and evaluate performance and scalability in various cooperative multi-agents environments.

**Assignment:** Multi-Agent Proximal Policy Optimization (MAPPO) is an advanced algorithm for multi-agent reinforcement learning. Social Influence Reward (SIR) is an innovative concept proposed by Jacques et al. (2019) that enhances the learning dynamics and adaptability of agents in multi-agent systems. The success of MAPPO in cooperative games, as demonstrated by Vinitzky et al. (2022), offers valuable insights for improving the performance of MAPPO.

Incorporating SIR into MAPPO can enhance agent cooperation and adaptability, which are crucial aspects of multi-agent systems. The success of MAPPO in cooperative games (Vinitzky et al., 2022) suggests that integrating SIR into MAPPO could lead to interesting improvements in multi-agent reinforcement learning.

The goal is to incorporate SIR into MAPPO and evaluate performance and scalability in various cooperative multi-agents environments.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- International Joint Conference on Artificial Intelligence [International Joint Conferences on Artificial Intelligence Organization (IJCAI Organization)]
- ICAART [SCITEVENT]

## Řešení problémů na vstupech se zakázanými vzory / Solving Problems on Pattern-Avoiding Inputs

**ID zadání / Topic ID:** 54

**Mentor:** Mgr. Michal Opler, Ph.D.

**Zadání:** Mnoho rozhodovacích problémů pracuje se sekvencí celých čísel, či množinou bodů v rovině (například problém třídění či geometrická verze problému obchodního cestujícího). V obou případech lze tyto vstupy přirozeně modelovat jako permutace, čili posloupnost  $n$  různých hodnot z množiny  $1, \dots, n$ . V kombinatorice se běžně studují vlastnosti permutací neobsahujících nějaký daný fixní vzor. Je proto přirozené ptát se, jak se změní složitost výpočetních problémů, pokud máme zaručeno, že jejich vstupy nebudou obsahovat nějaké zakázané vzory. Tyto zakázané vzory si můžeme představit jako bias v datech. V navrhovaném projektu bychom rádi prozkoumali různé takové rozhodovací problémy a vybrali si nějaký, o kterém ukážeme jak se změní jeho složitost na vstupech se zakázanými vzory.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- WG: International Workshop on Graph-Theoretic Concepts in Computer Science
- Information Processing Letters [Elsevier]
- arXiv [online]

## Masivně paralelní výpočet CRC na GPU / Massively Parallel CRC Calculation on GPU

**ID zadání / Topic ID:** 56

**Mentor:** Ing. Jiří Kašpar

**Mentor specialista / Mentor specialist:** doc. Ing. Ivan Šimeček, Ph.D.

**Zadání:** Nastudujte známé paralelní metody výpočtu CRC, vhodné z nich implementujte na GPU. Změřte výsledné rychlosti výpočtu a porovnejte je s dosud publikovanými výsledky.

**Assignment:** Read up on known parallel CRC algorithms. Select fitting algorithms and implement them on GPU. Measure speed of resulting implementations and compare them with known results.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- International Journal of Parallel, Emergent and Distributed Systems [Taylor & Francis Group]
- PPAM

## Complexity in Multi-Agent Systems / Complexity in Multi-Agent Systems

**ID zadání / Topic ID:** 58

**Mentor:** doc. RNDr. Dušan Knop, Ph.D.

**Zadání:** The aim of the project is to select some problem(s) (preferably in the area of Multi-Agent Systems) whose computational complexity is open. The complexity studied can vary from classical to parameterized or, e.g., special classes of inputs.

More students can apply for this topic and select their own problem.

**Assignment:** The aim of the project is to select some problem(s) (preferably in the area of Multi-Agent Systems) whose computational complexity is open. The complexity studied can vary from classical to parameterized or, e.g., special classes of inputs.

More students can apply for this topic and select their own problem.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- International Conference on Autonomous Agents and Multiagent Systems (AAMAS) [IFAAAMAS]
- International Workshop on Combinatorial Algorithms (IWOCA) [Springer LNCS]
- Workshop on Algorithms and Models for the Web Graph (WAW) [?]

# Moving Machine Learning Model from Simulated to Real Environments / Moving Machine Learning Model from Simulated to Real Environments

**ID zadání / Topic ID:** 59

**Mentor:** Ing Miroslav Čepek, Ph.D.

**Zadání:** Datalab owns robotic platforms for experimenting with autonomous driving - Deep Racer and Jet Racer. Both depends on camera for autonomous driving and the actions are the speed of the car and turning angle. Both also allow use of machine learning models to drive the car autonomously. Since training of machine learning models is highly impractical in real-world, the models are trained in simulated environment. The switch from simulated (ideal) environment to real-world is very challenging, as the real car has a lot of imprecisions, varying among cars and even single car can behave differently, depending with battery state.

The aim of this Vylet project is to explore possible improvements to techniques for transferring models trained in simulation into real hardware - like Sim2Real [1]. Additional aim is to simplify the transfer of the model between different cars of the same platform and even to different platforms.

[1] Allamaa, Jean Pierre, et al. "Sim2real for autonomous Vehicle control using executable digital twin." IFAC-PapersOnLine 55.24 (2022): 385-391.

**Assignment:** Datalab owns robotic platforms for experimenting with autonomous driving - Deep Racer and Jet Racer. Both depends on camera for autonomous driving and the actions are the speed of the car and turning angle. Both also allow use of machine learning models to drive the car autonomously. Since training of machine learning models is highly impractical in real-world, the models are trained in simulated environment. The switch from simulated (ideal) environment to real-world is very challenging, as the real car has a lot of imprecisions, varying among cars and even single car can behave differently, depending with battery state.

The aim of this Vylet project is to explore possible improvements to techniques for transferring models trained in simulation into real hardware - like Sim2Real [1]. Additional aim is to simplify the transfer of the model between different cars of the same platform and even to different platforms.

[1] Allamaa, Jean Pierre, et al. "Sim2real for autonomous Vehicle control using executable digital twin." IFAC-PapersOnLine 55.24 (2022): 385-391.

**Plánované uplatnění výstupů (časopisy, konference, jiné) / Planned output platforms (journals, conferences, etc.):**

- International Conference on Agents and Artificial Intelligence (ICAART)