# Návrh a ověření systému detekce anomálií založeného na strojovém učení v průmyslových řídicích systémech

Ing. Jan Vávra, Ph.D.

Teze disertační práce

Univerzita Tomáše Bati ve Zlíně

Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

Teze disertační práce

# Návrh a ověření systému detekce anomálií založeného na strojovém učení v průmyslových řídicích systémech

## Design and verification of anomaly detection system based on machine learning in industrial control systems

Autor: **Ing. Jan Vávra, Ph.D.**

Studijní program: Inženýrská informatika P3902
Studijní obor: Inženýrská informatika 3902V023

Školitel: doc. Ing. Luděk Lukáš, CSc.
Konzultant: doc. Ing. Martin Hromada, Ph.D.

Oponenti: prof. Mgr. Roman Jašek, Ph.D.
doc. Mgr. Ondrej Šuch, Ph.D.
doc. Ing. Petr Hrůza, Ph.D.

Zlín, prosinec 2020

Vydala **Univerzita Tomáše Bati ve Zlíně** v edici **Doctoral Thesis Summary.**
Publikace byla vydána v roce 2020

**Klíčová slova:** *kybernetická bezpečnost, strojové učení, umělá inteligence, detekce anomálií, průmyslový řídicí systém.*

**Key words:** *Cyber Security, Machine Learning, Artificial Intelligence, Anomaly Detection, Industrial Control System.*

Plná verze disertační práce je dostupná v Knihovně UTB ve Zlíně.

ISBN 978-80-7454-976-2

# ABSTRAKT

Technologie se staly nedílným prvkem současné společnosti. Současný přechod od průmyslové společnosti k informační společnosti je doprovázen implementací nových technologií do každé části lidské činnosti. Zvyšující se tlak na aplikaci informačních a komunikačních technologií v oblastech kritické infrastruktury a jejich řídicích systémů, zapříčiňuje vznik nových zranitelností. Tradiční přístupy pro zajištění bezpečnosti se stávají neefektivními. Z tohoto pohledu je využití umělé inteligence další evolučním krokem, který poskytuje robustní řešení i pro velmi rozsáhlé a komplexní systémy. Tato disertační práce je zaměřena na oblast výzkumu v rámci kybernetické bezpečnosti pro průmyslové řídicí systémy, které jsou četně využívány v kritické infrastruktuře. V rámci těchto systémů je každý výpadek možné definovat jako potencionální ohrožení základních služeb nutných pro chod společnosti. Hlavním jádrem disertační práce je vytvoření systému detekce anomálií, založeného na metodách strojového učení ve specifické oblasti kybernetické bezpečnosti průmyslových řídicích systémů. Zvláštní pozornost je poté věnována optimalizaci zvoleného řešení. Výsledný systém detekce anomálií je vytvořen s ohledem na jeho autonomní provoz a určitou míru interpretace kybernetických útoků.



# ABSTRACT:

Technology has become an integral part of contemporary society. The current transition from an industrial society to the information society is accompanied by the implementation of new technologies in every part of human activity. Increasing pressure to apply ICT in critical infrastructure and their control systems creates new vulnerabilities. Traditional safety approaches are becoming ineffective. From this perspective, the use of artificial intelligence is another evolutionary step that provides robust solutions for extensive and sophisticated systems. This dissertation focuses on the field of cybersecurity research for industrial control systems that are widely used in critical information infrastructure. Within these systems, each downtime can be defined as a potential threat to the core services needed to run the company. The main core of the thesis is to create an anomaly detection system based on machine learning methods in a specific area of cyber security of industrial control systems. Special attention is then paid to optimization. The resulting anomaly detection system is created concerning its autonomous operation and some degree of cyber attack interpretation.

# OBSAH

# 1. ÚVOD

Technologie v posledních několika dekádách zaznamenaly exponenciální růst v oblastech nasazení, efektivity a funkcionality. Systémy založené na zmíněných technologiích využívají automatizace, digitalizace, robotizace a jsou často vzájemně propojeny s využitím vzdáleného řízení. Mluví se o revoluci v podobě tzv. „Průmyslu 4.0“, který v budoucnu ovlivní většinu aspektů lidské společnosti. Frank [1] ve své publikaci prezentoval základní atributy, ve kterých je Průmysl 4.0 uplatňován. Jedná se o integraci, řízení energií, vystopovatelnost (tzv. „tracebility“), automatizaci, virtualizaci a flexibilitu pomocí technologií jako je internet věcí (internet of thinks – IoT), „cloud computing“. „big data“ a analytika. Právě dvě posledně jmenované oblasti („big data“ a analytika) jsou podle řady autorů považovány za největší motory průmyslové revoluce 4.0. [1] Do této skupiny řadíme techniky dolování dat („data mining“) a strojového učení („machine learning“).

Automatizované a autonomní nahrazení klasické lidské činnosti je cílem popisované revoluce. Z důvodu zvýšení efektivity jsou komunikačně propojovány i technologické celky, spadající do oblasti kritické infrastruktury (KI). Z tohoto důvodu jsou systémy KI zatíženy značným tlakem z pohledu relativně nových hrozeb, jakými jsou například kybernetické útoky, které se z historického hlediska stávají zásadní hrozbou pro dnešní a budoucí společnost. Zvláště ohrožené jsou řídicí průmyslové systémy (Industrial Control System – ICS), jelikož představují skupinu systémů přímo ovlivňujících fyzický svět. Zásadní přehodnocení ICS kybernetické bezpečnosti bylo zapříčiněno působením počítačového červa Stuxnet v roce 2010. [2]

Předložená disertační práce je zaměřena na detekci anomálií v oblasti průmyslových řídicích systémů z pohledu kybernetické bezpečnosti. Důraz je kladen na respektování specifik ICS systémů. Ty jsou často rozdílné od požadavků běžně využívané výpočetní techniky. Pro detekci anomálií jsou využity metody, techniky a algoritmy vztahující se ke strojovému učení a dolování dat. Každý ze zvolených postupů je podrobně analyzován v následujících kapitolách. V závislosti na vybraném nastavení a vytvořené hypotéze jsou vyhotoveny experimenty, které hodnotí zvolená řešení.

# 2. SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

Důležitost ochrany před kybernetickými útoky je primárně svázána s důležitostí chráněného aktiva. V tomto smyslu slova, lze konstatovat, že bezpečnost ICS systémů je jednou z nejkritičtějších oblastí bezpečnosti z důvodu kritické infrastruktury, ve které jsou ICS systémy využity. S příchodem čtvrté (digitální) průmyslové revoluce se řešená problematika kybernetické bezpečnosti stala ještě více významnou. Z tohoto důvodu je kybernetická bezpečnost ICS systémů velmi důležitou oblastí výzkumu. Tato disertační práce je zaměřena na detekci kybernetických útoků pomocí rozpoznávání anomálií komunikačního provozu za pomocí algoritmů strojového učení.

Existují dvě hlavní skupiny detekčních metod pro kybernetické útoky. Toto rozdělení podporují publikace [3], [4]. První základní skupina detekce, založená na pravidlech, je také známa pod názvem detekce signatur (Signature detection). Tato oblast detekce má však zásadní slabinu ve způsobu činnosti. Každá signatura přesně odpovídá jednomu kybernetickému útoku, který musel být nejprve zaregistrován, analyzován a na základě této analýzy může být vydána signatura, která je distribuovaná do databází signatur. Z tohoto důvodu je každý IDS systém účinný do té míry, do jaké má aktuální a kvalitní databázi signatur. Proces identifikace a analýzy kybernetických útoků je v mnoha případech značně náročný.

Do druhé skupiny lze zařadit oblast detekce anomálií, někdy také nazývanou jako detekce odlehlých hodnot nebo odchylek, která je určitým nástrojem pro oddělení normálního nezávadného chování od často výjimečného až anomálního chování, které často slouží jako příznak kybernetického útoku. Tato oblast detekce je do značné míry založena na metodách umělé inteligence, respektive na strojovém učení.

Tato skupina algoritmů strojového učení vychází z oblastí učení s učitelem a učení bez učitele. Podle Obr. 1 vypadá tato skupina algoritmu blíže učení s učitelem nežli učení bez učitele. Avšak jsou zde určité rozdíly. Vstupní data zastupují jenom jednu třídu. V oblasti kybernetické bezpečnosti jsou využívána data reprezentující normální a tím pádem i bezpečný provoz. Tato data jsou upravena a rozdělena do skupin (validační dataset, trénovací dataset a testovací dataset). Vytvořený model musí být „natrénovaný“ a validovaný na datech, která se vztahují jen k jedné třídě. Tato data obsahují jen normální a bezproblémový provoz sledovaného systému bez přítomnosti anomálií (kybernetických útoků). Kybernetické útoky jsou obsaženy jen v rámci testovacího datasetu z důvodu evaluace vytvořeného modelu. Výsledná data jsou poté rozdělena pomocí prediktivního modelu do dvou skupin. První skupina

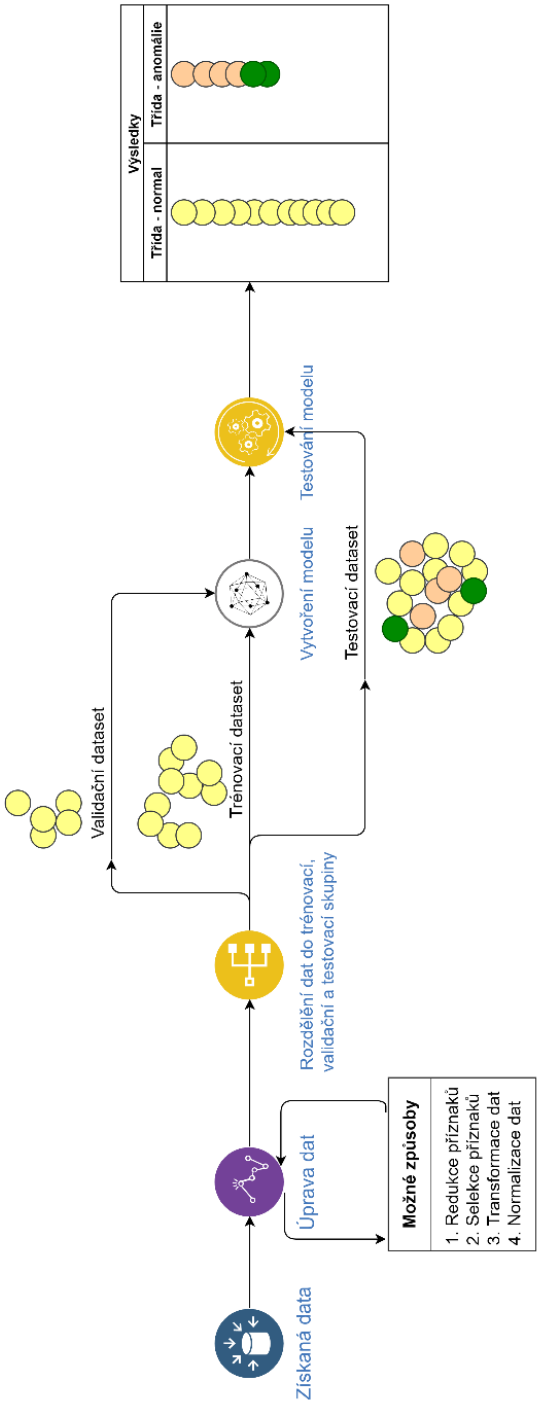
představuje data, která prediktivní model definuje jako normální provoz chráněného systému. Druhou skupinou je v podstatě vše ostatní. Tedy data, která se odlišují od dat využitých k trénování. Výhodou tohoto řešení je detekce neznámého chování ve sledovaných datech.

Využití metod strojového učení pro ochranu ICS před kybernetickými útoky je poměrně nová a dynamická oblast kybernetické bezpečnosti. Tato specifická oblast bezpečnosti se stala předmětem zájmu po kybernetickém incidentu v Íránu (2010), kde počítačový červ Stuxnet destabilizoval jadernou elektrárnu v Búšehru. Tento milník vedl k zásadnímu přehodnocení kybernetické bezpečnosti v oblasti ICS. Následný vývoj začalo ve svých publikacích reflektovat značné množství autorů, kteří hledali nové způsoby ochrany ICS před kybernetickými útoky. V posledních letech se stále více autorů zaměřuje na kybernetickou ochranu ICS prostřednictvím technik pro detekci anomálií. Sokolov [5] ve své publikaci uvedl základní rozdělení, výhody a nevýhody aplikace algoritmů strojového učení v prostředí ICS. Konstatoval nevhodnost použití systému detekce anomálií, založeném na lineárních modelech z důvodu nízké efektivity detekce anomálií. Dále vyzdvihl lepší detekční schopnosti skupiny technik založených na stromových strukturách (např. Random Forest). Výsledky však podle autora ukazují na náchylnost k „přeučení“ vytvořeného modelu, a tudíž i horším generalizačním vlastnostem tohoto řešení. Poslední zkoumanou skupinou byly podle autora neuronové sítě. Ty vykazovaly nejpřesnější výsledky, avšak za cenu vyšší výpočetní náročnosti, zvláště pak při velkém počtu vstupních dat.

Mnozí autoři popisují oblast detekce anomálií v rozsáhlém počtu publikací pro různé oblasti lidské činnosti. Avšak jen omezené procento z nich je zaměřeno na oblast kybernetické bezpečnosti pro ICS systémy. Značné množství autorů představuje slibná řešení, která však neberou v potaz aspekty a kritéria pro využití těchto detekčních systémů v prostředí ICS. Vyvstává proto řada otázek, na které je nutné najít dostatečnou odpověď. Jednou ze zásadních otázek pro nasazení metod strojového učení je výpočetní, a tudíž i časová náročnost detekčních modelů pro systém ICS. V řadě publikací není zohledněna problematika falešných poplachů a jejich minimalizace z důvodu ochrany dostupnosti služeb ICS a zajištění jeho kontinuálního chodu. Toto kritérium, skýtá značné riziko pro systémy ICS. Základní nedostatky současného řešení se dají shrnout do následujících bodů:

- Nízká míra přesnosti v rámci identifikace kybernetických útoků,
- Vysoká míra falešných poplachů,
- Vysoká výpočetní náročnost,

- Interpretace kybernetického útoku.




Obr. 1: Detekce anomálií založená na kombinaci strojového učení. [vlastní zdroj]

# 3. CÍLE DISERTAČNÍ PRÁCE

Hlavním cílem dizertační práce je:

**Konceptuální návrh a ověření systému detekce anomálií z pohledu kybernetické bezpečnosti, založeného na strojovém učení, v průmyslových řídicích systémech.**

K dosažení hlavního cíle bude nutné splnit tyto dílčí cíle:

- vymezení postupu identifikace kybernetických útoků pro systémy ICS,
- výběr, úprava a analýza vybraných ICS datasetů a jejich parametrů, které budou využity pro detekci anomálií,
- identifikace a analýza algoritmů strojového učení vhodných pro oblast detekce anomálií,
- využití optimalizačních technik pro zvýšení detekčních schopností zvoleného řešení,
- zhodnocení možnosti interpretace detekovaných anomálií,
- vytvoření algoritmu pro detekci anomálií, založeném na strojovém učení,
- ověřování, testování a hodnocení navrženého řešení.

# 4. ZVOLENÉ METODY ZPRACOVÁNÍ

Pro úspěšné řešení cílů dizertační práce bylo využito následujících metod vědecké práce (metoda analýzy, metoda syntézy, metoda modelování, metoda komparace, metoda experimentu, metoda matematické statistiky, metoda indukce).

- **Metoda analýzy** – jedná se o obecnou vědeckou metodu, která je založena na rozkladu zkoumaného jevu na dílčí části, které jsou dále vyhodnoceny za účelem odhalení jejich podstaty. Pro účely disertační práce je tato metoda využita k porozumění technik, dat a postupů, čehož bylo následně využito k jejich výběru.
- **Metoda syntézy** – tato vědecká metoda je založena na spojení dílčích částí do jednoho celku při sledování souvislostí, vazeb a vlastností mezi jednotlivými prvky jevu. V disertační práci je využita pro tvorbu systému detekce anomálií. K tomu je použito řady dílčích postupů, technik a algoritmů.
- **Metoda modelování** – metodu modelování můžeme definovat jako experimentální proces, jehož výsledkem je vytvoření abstraktního modelu, který je využit pro detekci komunikačních anomálií. V souvislosti s cíli dizertační práce je využito metody modelování k vytvoření prediktivního modelu pro klasifikaci zvoleného datasetu.
- **Metoda komparace –** jedná se o základní metodu pro porovnání dvou a více objektů v jednotném prostředí (stejné podmínky), popřípadě pro vyhodnocení dvou a více prostředí (rozdílné podmínky) pro jeden objekt. V rámci dizertační práce je použito metody komparace pro porovnání účinnosti jednotlivých metod detekce. K tomu je využito metod matematické statistiky.
- **Metoda experimentu** – tato empirická metoda je zaměřena na testování a ověřování pravdivosti vytvořených hypotéz za stanovených podmínek. Cílem je verifikovat neboli potvrdit nebo falzifikovat neboli vyvrátit platnost hypotézy. Metoda experimentu je jednou z důležitých metod nutných pro naplnění cílů dizertační práce. Tato metoda je využita k ověření předpokladů v oblasti detekce anomálií.
- **Metody matematické statistiky** – jedná se o exaktní disciplínu, která je založena na analýze empirických dat. Využití těchto metod umožňuje analyzovat data za účelem přesné specifikace jevů a jejich vztahů. Metody matematické statistiky jsou v rámci dizertační práce využity pro analýzu datových setů a v oblasti vývoje vhodných postupů při identifikaci anomálií.
- **Metoda indukce** – tato vědecká metoda je založena na generalizaci sledovaných jevů. Výsledek indukce vypovídá o podstatě zákonitostí ve sledovaném jevu. Výstupem této vědecké metody je hypotéza, která je buď potvrzena, nebo vyvrácena. Tato vědecká metoda je využita ke tvorbě hypotéz v rámci oblasti detekce anomálií.

# 5. TEORETICKÝ RÁMEC

Účelem této kapitoly je vytvořit teoretický základ použitých technik a algoritmů za účelem objasnění zvolených postupů využitých pro naplnění vytyčených cílů disertační práce. Tato kapitola je rozdělena do osmi podsekcí, které reflektují základní úkony, které byly nezbytné pro naplnění cíle disertační práce.

## 5.1 Úprava dat

Úprava a výběr vstupních dat je jednou z velmi důležitých částí pro efektivní využití algoritmu strojového učení. Často však jsou tato data ukládána v rozdílných formátech. Proto je nutná jejich transformace do podoby, která bude vyhovovat příslušným algoritmům. Je vhodné poznamenat, že obecně strojové učení využívá numerické hodnoty pro trénování modelů. V rámci disertační práce je pro numerické hodnoty využita **transformace dat na jednotné měřítko**. Druhá transformace se týká problematiky **chybějících hodnot**. Pro transformaci nominálních dat bylo využito transformace **„one-hot encoder“**. Využitá data jsou dále nazývána jako dataset.

## 5.2 Výběr atributů

Tato oblast také známá jako výběr rysů, je základním úkonem k vytvoření efektivního modelu v oblasti strojového učení. Jedná se o výběr nejvhodnější podmnožiny dat z celého datasetu, který může obsahovat stovky až tisíce atributů. Hlavní myšlenou této oblasti je tvorba redukovaného datasetu při zachování jeho informační hodnoty. Z důvodu redukce dimenze využívaných datasetů bylo nutné vybrat řešení, které efektivně sníží dimenzi datasetu a umožní interpretaci výsledků. Jako vhodné řešení tohoto problému byla vybrána metoda analýzy hlavních komponent (**Principal Component Analysis – PCA**). Základní myšlenkou PCA je redukovat počet atributů, a přitom zachovat jejich původní informační hodnotu.[6]

## 5.3 Analýza využívaných datasetů

V této podkapitole jsou analyzovány tři datasety, které byly využity pro tvorbu a testování systému detekce anomálií. **První z datasetů** byl představen v publikaci [7]. Dataset sestává z několikahodinového záznamu síťové komunikaci ICS systému. Tato data byla vygenerována pomocí řady simulací. Výsledný systém představuje elektrickou distribuční síť se zdrojem o 12 000 V. Vytvořený SCADA sandbox je složen z několika MTU a RTU, které komunikují prostřednictvím Modbus komunikačního protokolu.[7] Tento dataset obsahuje čtyři kybernetické útoky (CA1_1 až CA1_4). **Druhý dataset** byl vygenerován na univerzitě: „University of

Technology and Design“ v Singapuru.[8] Jedná se o reálný systém určený pro čištění odpadních vod. Tento dataset obsahuje šest kybernetických útoků pro tvorbu systému detekce anomálií (CA2_1 až CA2_6) a další tří pro ověření výsledků (CA2_7 až CA2_9). **Třetí dataset** byl vytvořen na „Mississippi State University“. Svým účelem představuje plynovod viz. [9]. Tento systém obsahuje akční členy jako je například pumpa nebo solenoid pro kontrolu tlaku v potrubí. Tento dataset obsahuje šest kybernetických útoků pro tvorbu systému detekce anomálií (CA3_1 až CA3_6) a další tří pro ověření výsledků (CA3_7 až CA3_9).

## 5.4 Popis zvolených algoritmů

V rámci této podkapitoly jsou popsány čtyři algoritmy strojového učení, kterých byly využito v disertační práci. Všechny algoritmy spadají do oblasti strojového učení využívající kombinace principů učení s učitelem a učení bez učitele. **Neuronové sítě (ANN)** vycházejí svojí filozofií z fungování řídicího centra nervové soustavy u biologických organismů. Jako druhý algoritmus byl vybrán **One-class Support Vector Machines(OCSVM)**. Základem **OCSVM** je vytvoření nadroviny, která odděluje data. Tuto nadrovinu je nutné maximalizovat, tedy oddělit od sebe dvě skupiny dat. Třetí algoritmus strojového učení byl vybrán **„Isolation Forest“ (IF).** IF vychází ze dvou předpokladů. První z předpokladů vychází ze skutečnosti, že anomálie jsou zastoupeny v datech velmi zřídka. Druhý předpoklad vychází z rozdílnosti hodnoty mezi atributem normálního záznamu a atributem anomálního záznamu. Posledním algoritmem strojového učení byl zvolen **LSTM (Long short-term memory)**. Jedná se o algoritmus strojového učení, který spadá do podskupiny rekurentních neuronových sítí. Tyto algoritmy pracují se sekvenčními daty, kde významnou roli hrají nejenom hodnoty jednotlivých atributů, ale také jejich uspořádání. [10]

## 5.5 Optimalizace

Pro nalezení optimálního řešení v rámci systému detekce anomálií byly využity tři optimalizační algoritmy. Tyto algoritmy jsou využity pro nastavení algoritmů strojového učení (hyperparametry) v závislosti na cílové funkci **(CF)**. Jako první optimalizační algoritmus byl vybrán **Random Search (RS)**. RS využívá kombinatoriky, kde sestavuje množinu všech možných kombinací využívaných hyperparametrů. Z této množiny poté náhodně vybírá zástupce, kteří jsou následně otestováni. Druhý zvolený algoritmus je **Genetický Algoritmus (GA)**. Jedná se o robustní vyhledávací algoritmy založené na heuristice, která vychází z Darwinovy evoluční teorie. Kde jenom atributy, vlastnosti těch nejlepších jedinců přecházejí na další generace. GA proto konvergují do jednoho optimálního řešení (jednotlivce),

který je vybrán simulovanou evolucí. Jako poslední algoritmus byl zvolen **Tree-structured Parzen Estimator (TPE).** Jedná se o optimalizační algoritmus založený na Bayesovské optimalizaci, která využívá Gaussova procesu. TPE popsali autoři Bergstra et al. (2011), ve své publikaci [11].

## 5.6 Multikriteriální hodnocení

Základní myšlenkou a úkolem **multikriteriálního hodnocení (MH)** je v tomto případě výběr jedné varianty z množiny možností na základě vybraných kritérií. V rámci disertační práce je využito multikriteriální hodnocení pro definici **CF**, která je nezbytná pro optimalizační algoritmy. **CF** je v tomto případě definována pomocí pěti metrik (viz. kapitola 5.7). V rámci řešené problematiky byla využita metoda multikriteriálního hodnocení **TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution)**. Mezi klady této metody patří nízká výpočetní náročnost a konzistentnost metody. Tato metoda bere také v potaz nejenom nejlepší možné výsledky, ale také ty nejhorší, což umožnuje negovat špatné výsledky v jednom kritériu za dobré výsledky v jiném.

## 5.7 Hodnocení výsledků

Tato kapitola popisuje metodiku hodnocení klasifikačních modelů, která je využívána v rámci odborné komunity. Každý z využitých algoritmů strojového učení je ohodnocen pomocí pěti metrik. Tyto metriky spolu s metodou TOPSIS slouží pro výpočet **CF**. Výčet těchto metrik je následující:

- $\mathbf{M_{MCC}}$ – (Matthews Correlation Coefficient). Výsledná hodnota metriky vyjadřuje všechny aspekty konfúzní matice. Avšak není zatížena nedostatky pří využití nevyváženého datasetu jako tomu je u metriky “Accuracy”.
- $\mathbf{M_{F1}}$ – tato metrika (F1 skóre) je dosti využívána pro popis výsledků konfúzní matice. Oproti metrice “Accuracy” není tolik náchylná na výskyt nevyvážených tříd v datasetu.
- $\mathbf{M_{Prec}}$ – tato metrika (precision) vypočítává pravděpodobnost pozitivní klasifikace. Vyjadřuje poměr správně pozitivně identifikovaných prvků ke všem prvkům, které jsou označeny jako pozitivní.
- $\mathbf{M_{FPR}}$ – tato **nejvýznamnější metrika** (False positive rate (FPR)) vyjadřuje případ, kdy pozitivní třídy jsou identifikovány jako falešné, tedy jedná se o případ, kdy normální a nezávadná komunikace na počítačové síti je vyhodnocena jako nebezpečná.
- **Čas** – toto kritérium vyjadřuje čas potřebný k predikci a klasifikaci testovacího datasetu prostřednictvím prediktivního modelu.

# 6. HLAVNÍ VÝSLEDKY PRÁCE

V rámci této kapitoly jsou prezentovány jednotlivé výsledky výzkumu. Kapitola je rozdělena do pěti podkapitol. Každá z nich přispěla k naplnění vytyčených cílů disertační práce. Výčet podkapitol je následující: metoda výběru a identifikace kritických kybernetických útoků vůči systému ICS, úprava datasetů před jejich využitím pro algoritmy strojového učení, optimalizace algoritmů strojového učení, ověření a interpretace systému detekce anomálií.

## 6.1 Identifikace současných hrozeb a zranitelností ICS v kybernetickém prostoru

Pro naplnění tohoto cíle byl využit Shodan nástroj pro detekci a identifikaci internetově připojených zařízení a ICS-CERT databáze zranitelností, která poskytuje aktuální seznam ICS zranitelností. Hlavním předpokladem této části výzkumu byla časová prodleva mezi zveřejněním zranitelnosti v databázi ICS-CERT a reálnou aktualizací systému z důvodu odstranění zranitelnosti. Tato zranitelnost byla umocněna faktem, že samotné aktualizace jsou pro ICS, a tudíž i pro SCADA systémy velmi kritické, a proto je nelze provádět na denní bázi z důvodu jejich testování. V roce 2016 bylo shromážděno 974 zranitelných ICS systémů. Ty i přes doporučení organizací ICS-CERT na jejich izolaci od internetového připojení byly reálně dosažitelné právě prostřednictvím internetového připojení. Nejvíce zasaženou zemí byly Spojené státy americké, kde bylo 487 zranitelných systémů. Na druhém místě bylo Španělsko se 75 zranitelnými systémy a na třetím místě Kanada s 59 zranitelnými systémy. Z výsledků vyplývá, že na celou Evropu připadá 291 zranitelných systémů, což nepřesahuje celkový počet postižených systémů v USA. Lze také konstatovat, že téměř 50 % všech zasažených systémů bylo zasaženo zranitelností ICSA-16-026-02.

## 6.2 Úprava datasetů

V této podkapitole jsou prezentovány výsledky řady experimentů. Na jejich základě byly zvoleny optimální postupy a techniky pro úpravu datasetů z pohledu kybernetické bezpečnosti ICS. Byly využity techniky pro náhradu prázdných hodnot a technik pro změnu měřítka. Pro tvorbu a evaluaci dílčích konfigurací technik byl zvolen dataset 1 se čtyřmi kybernetickými útoky. Pro každý ze čtyř algoritmů strojového učení a následně pro každý ze čtyř kybernetických útoků bylo vytvořeno devět kombinací technik pro úpravu dat (data1 až data9). Tři techniky pro náhradu chybějící hodnoty a tři techniky pro změnu měřítka. Tyto kombinace jsou vytvořeny v následujícím pořadí: **data1** – aritmetický průměr; normalizace <0,1>, **data2** –

aritmetický průměr; normalizace <-1,1>, **data3** – aritmetický průměr; standardizace, **data4** – medián; normalizace <0,1>, **data5** – medián; normalizace <-1,1>, **data6** – medián; standardizace, **data7** – náhrada konstantou; normalizace <0,1>, **data8** – náhrada konstantou; normalizace <-1,1>, **data9** – náhrada konstantou; standardizace.

Bylo vytvořeno 900 prediktivních modelů pro každý dílčí algoritmus strojového učení. Pro evaluaci řešení byly použity následující metriky: $M_{F1}$, $M_{MCC}$, $M_{Prec}$, $M_{FPR}$ a Čas. Z důvodu porovnání všech variant zpracování dat v rámci každého algoritmu strojového učení bylo využito neparametrického testu – Friedmanův test [12]. V rámci tohoto testu je potvrzena nebo vyvrácena nulová hypotéza pomocí p-hodnoty. Pro zamítnutí nebo přijetí nulové hypotézy je uvažováno s hodnotou 5 %. Pro zhodnocení statistické významnosti výsledků Friedmanova testu je využito Nemenyiho testu pro definici kritické vzdálenosti. Ta určuje významné statistické rozdíly mezi daty. Souhrnné výsledky jsou prezentovány v Tab. 1. Každá kombinace byla ohodnocena podle Friedmanova testu.

Tab. 1 – Souhrnné výsledky experimentu – úprava datasetu. [vlastní zdroj]

| **Kombinace technik/ algoritmy** | **data 1** | **data 2** | **data 3** | **data 4** | **data 5** | **data 6** | **data 7** | **data 8** | **data 9** |
|---|---|---|---|---|---|---|---|---|---|
| Neuronová síť | 5 | 6 | 1 | 4 | 3 | 2 | 4 | 8 | 7 |
| | 1 | 5 | 8 | 2 | 4 | 7 | 3 | 3 | 6 |
| | 1 | 2 | 6 | 3 | 5 | 8 | 4 | 4 | 7 |
| | 1 | 5 | 7 | 3 | 2 | 8 | 4 | 5 | 6 |
| Suma | 8 | 18 | 22 | 12 | 14 | 25 | 15 | 20 | 26 |
| LSTM | 1 | 4 | 7 | 2 | 5 | 6 | 3 | 9 | 8 |
| | 1 | 4 | 7 | 2 | 5 | 8 | 3 | 6 | 9 |
| | 1 | 6 | 9 | 2 | 4 | 8 | 3 | 5 | 7 |
| | 1 | 4 | 9 | 2 | 6 | 8 | 3 | 5 | 7 |
| Suma | 4 | 18 | 32 | 8 | 20 | 30 | 12 | 25 | 31 |
| IF | 8 | 3 | 2 | 7 | 5 | 6 | 9 | 4 | 1 |
| | 8 | 4 | 1 | 7 | 6 | 2 | 9 | 5 | 3 |
| | 4 | 8 | 3 | 2 | 9 | 5 | 6 | 7 | 1 |
| | 6 | 4 | 1 | 5 | 7 | 2 | 8 | 9 | 3 |
| Suma | 26 | 19 | 7 | 21 | 27 | 15 | 32 | 25 | 8 |

Algoritmus OCSVM byl kvůli diametrálně horším výsledkům vyňat z množiny testovaných algoritmů strojového učení. Ze souhrnných výsledků lze vždy vyčlenit vhodné řešení pro každého zástupce z algoritmů strojového učení. V tomto případě nižší výsledné číslo značí lepší výsledky. V případě neuronové sítě se jedná o

kombinaci data1. V případě LSTM se jedná o kombinaci data1 a v případě algoritmu IF se jedná o kombinaci technik spadajících pod označení data3.Tento experiment byl významný pro další postup disertační práce. Další experimenty budou vycházet z jeho výsledků při úpravě datasetů.

## 6.3 Postup nastavení a ohodnocení jednotlivých algoritmů strojového učení pomocí optimalizačních technik

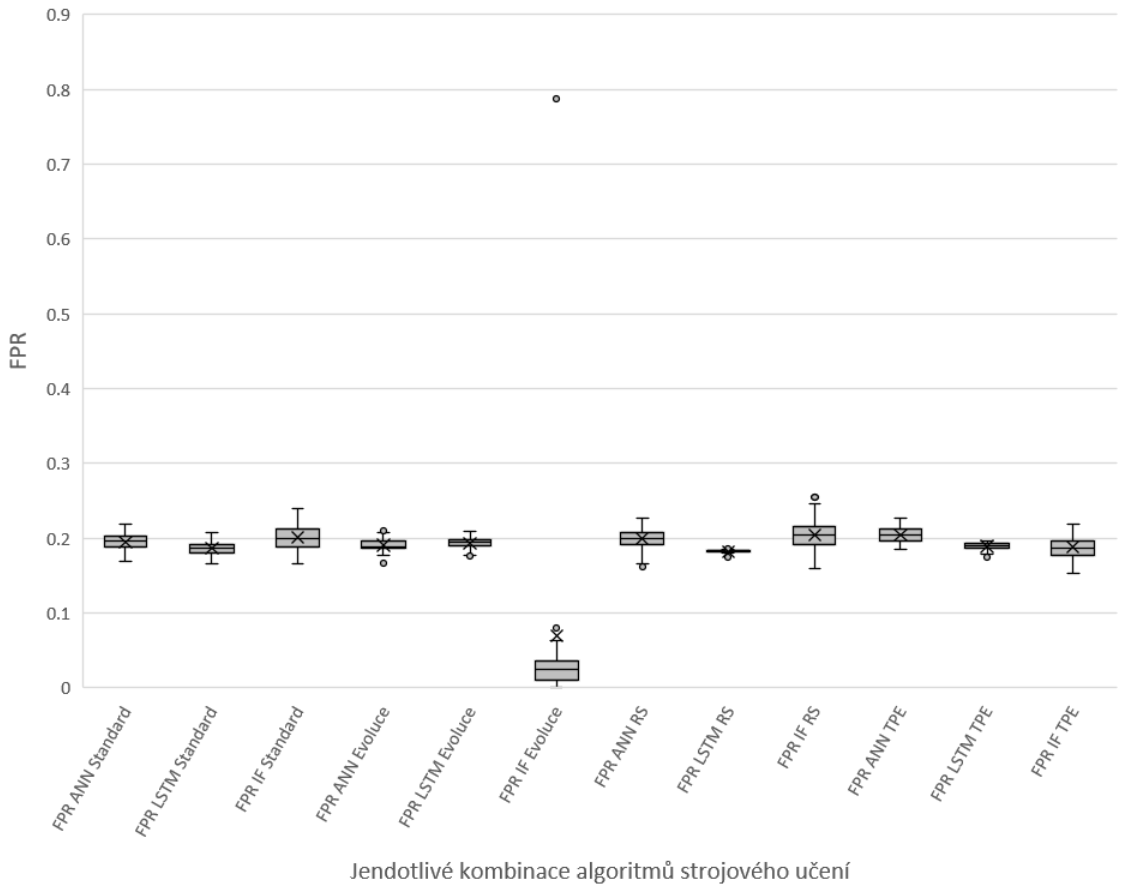
Tato podkapitola disertační práce je zaměřena na popsání procesu a diskusi výsledků optimalizace hyperparametrů pro vybrané algoritmy strojového učení. Na základě výsledků optimalizačních algoritmů byla vybrána optimální nastavení (hyperparametrů) jednotlivých algoritmů strojového učení pro vybrané datasety. Výsledky z provedených experimentů jsou srovnány prostřednictvím pěti metrik $M_{F1}$, $M_{MCC}$, $M_{Prec}$, $M_{FPR}$ a Čas pomocí Friedmanova testu viz. Tab. 2. V rámci každého grafu je diskutováno 12 hodnot odpovídajících jednotlivým kombinacím od data1 do data12. Kde data1, data2 a data3 představují základní nastavení algoritmů v pořadí neuronová síť, LSTM, IF. Data4 až data 6 reprezentují algoritmy neuronová síť, LSTM, IF při nastavení hyperparametrů pomocí evolučního algoritmu. Data7 až data9 představují algoritmy neuronová síť, LSTM, IF při nastavení hyperparametrů pomocí RS. Data10 až data 12 představují algoritmy neuronová síť, LSTM, IF při nastavení hyperparametrů pomocí TPE. Druhý experiment byl výhradně zaměřen na metriku $M_{FPR}$, která je nejdůležitějším indikátorem pro ICS systémy.

Tab. 2 – Komparace jednotlivých algoritmů podle pořadí v rámci Friedmanova testu. [vlastní zdroj]

| **Zastoupené algoritmy/ datasety** | **data1** | **data2** | **data3** | **data4** | **data5** | **data6** | **data7** | **data8** | **data9** | **data10** | **data11** | **data12** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dataset 1 | 4 | 2 | 10 | 3 | 5 | 11 | 1 | 6 | 7 | 1 | 8 | 9 |
| | 5 | 2 | 10 | 4 | 9 | 7 | 3 | 7 | 6 | 1 | 11 | 8 |
| | 6 | 5 | 11 | 2 | 4 | 7 | 3 | 7 | 10 | 1 | 9 | 8 |
| | 3 | 1 | 9 | 5 | 6 | 12 | 4 | 8 | 11 | 2 | 7 | 10 |
| Dataset 2 | 2 | 3 | 3 | 8 | 7 | 2 | 1 | 6 | 5 | 4 | 9 | 2 |
| | 1 | 7 | 10 | 8 | 5 | 4 | 3 | 11 | 9 | 6 | 2 | 12 |
| | 3 | 2 | 4 | 6 | 6 | 9 | 9 | 1 | 5 | 10 | 7 | 8 |
| | 1 | 4 | 8 | 4 | 6 | 7 | 2 | 11 | 9 | 10 | 3 | 5 |
| | 10 | 4 | 2 | 8 | 9 | 5 | 6 | 12 | 1 | 7 | 11 | 3 |
| | 1 | 5 | 9 | 4 | 6 | 6 | 3 | 7 | 10 | 9 | 2 | 8 |
| Dataset 3 | 5 | 10 | 2 | 1 | 8 | 6 | 3 | 9 | 11 | 4 | 1 | 7 |
| | 9 | 10 | 2 | 5 | 11 | 4 | 7 | 3 | 12 | 1 | 8 | 6 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 6 | 4 | 1 | 7 | 4 | 4 | 9 | 6 | 2 | 8 | 3 | 5 |
| | 6 | 9 | 10 | 1 | 7 | 7 | 5 | 3 | 11 | 2 | 4 | 8 |
| | 1 | 6 | 8 | 3 | 7 | 11 | 2 | 9 | 10 | 5 | 4 | 12 |
| | 10 | 7 | 3 | 1 | 6 | 8 | 4 | 11 | 9 | 2 | 5 | 7 |
| Suma | 73 | 81 | 102 | 70 | 106 | 110 | 65 | 117 | 128 | 73 | 94 | 118 |

Mezi nejlepší zástupce z výsledné tabulky lze zařadit následující algoritmy: neuronová síť nastavená podle RS a neuronová síť nastavená podle evolučního algoritmu. Tyto výsledky se však nedají považovat za jednoznačné, a to především proto, že v rámci prezentovaných grafů jsou zobrazené algoritmy porovnávány podle pěti metrik, přičemž v tomto případě jsou všechny metriky stejně významné. Tedy žádná metrika nevybočuje svojí významností nad ostatní. Proto byl uskutečněn druhý experiment, kde výsledky byly prezentovány pomocí krabicového grafu viz. Obr.2. V tomto obrázku jsou porovnáni všichni zástupci vybraných algoritmů. Tento ilustrační obrázek je jedním z výsledků pro kybernetický útok CA2_3 ve druhém datasetu. Ze souhrnných výsledků vyplývá následující. Pro dataset 2 a 3 je nejlepší možností algoritmus IF nastavený podle evolučního algoritmu.




Obr. 2: Porovnání metriky $M_{FPR}$ pro jednotlivé kombinace algoritmu strojového učení a optimalizačních algoritmů pro kybernetický útok – CA2_3. [vlastní zdroj]

Výsledky ukazují prakticky nulové hodnoty metrik $M_{FPR}$ (nejlepší možný výsledek). V případě datasetu 1 se ukázalo, že využití optimalizačního algoritmu TPE pro nastavení hyperparametrů neuronové sítě vykazuje nejlepší výsledky.

## 6.4 Ověření systému detekce anomálií

Tato sekce je zaměřena na ověření vytvořeného systému detekce. Pro tento účel bylo vybráno šest kybernetických útoků. Tři z datasetu 2 (CA2_7, CA2_8, CA2_9) a tři z datasetu 3 (CA3_7, CA3_8, CA3_9). Byly provedeny dva typy experimentů, stejně jako v případě podkapitoly 6.3. Bylo využito následujících kombinací: data1, data2 a data3 reprezentují algoritmy neuronová síť, LSTM, IF při nastavení hyperparametrů pomocí evolučního algoritmu. Data4 až data6 představují algoritmy neuronová síť, LSTM, IF při nastavení hyperparametrů pomocí RS. Data7 až data9 představují algoritmy neuronová síť, LSTM, IF při nastavení hyperparametrů pomocí TPE. Výsledky jsou uvedeny v Tab. 3.

Tab. 3 – Komparace jednotlivých algoritmů podle pořadí v rámci Friedmanova testu (ověření výsledků). [vlastní zdroj]

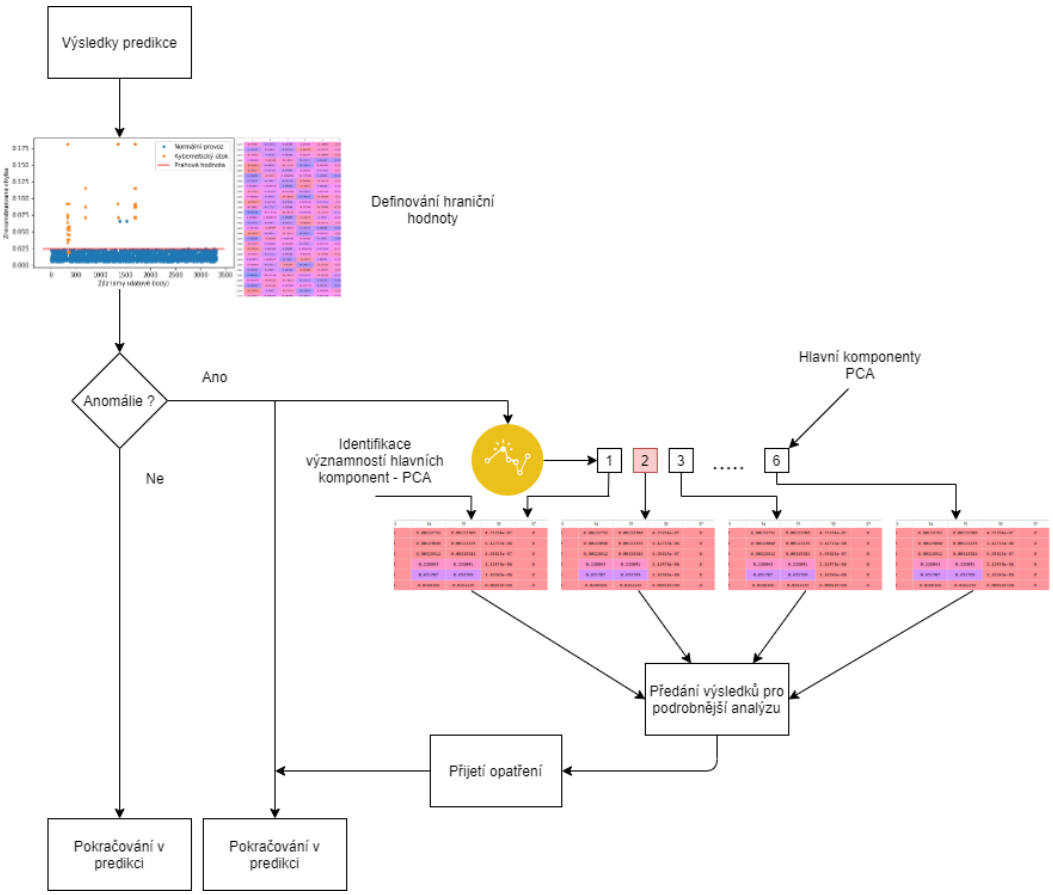
| **Zastoupené algoritmy/ datasety** | **data1** | **data2** | **data3** | **data4** | **data5** | **data6** | **data7** | **data8** | **data9** |
|---|---|---|---|---|---|---|---|---|---|
| Dataset 2 | 1 | 3 | 2 | 4 | 7 | 5 | 6 | 2 | 5 |
| | 3 | 4 | 1 | 2 | 5 | 6 | 7 | 2 | 6 |
| | 7 | 2 | 1 | 3 | 1 | 6 | 4 | 5 | 8 |
| Dataset 3 | 1 | 5 | 3 | 2 | 6 | 7 | 4 | 8 | 5 |
| | 8 | 2 | 2 | 6 | 5 | 4 | 7 | 3 | 1 |
| | 4 | 3 | 1 | 6 | 9 | 8 | 7 | 5 | 2 |
| Suma | 24 | 19 | 10 | 23 | 33 | 36 | 35 | 25 | 27 |

Mezi nejlepší zástupce z výsledné tabulky lze zařadit následující algoritmy: algoritmus IF nastavený prostřednictvím evolučního algoritmu (data3) – skóre 10 a algoritmus LSTM nastavený podle evolučního algoritmu (data2) – skóre 19. Tyto výsledky jsou poměrně zajímavé, zvláště pak při srovnání s výsledky v Tab. 3. Algoritmus IF, který má vybrané hyperparametry podle evolučního algoritmu zde ukazuje své dominantní umístění v porovnání se zbylými variantami algoritmů strojového učení.

V případě metriky $M_{FPR}$ lze identifikovat nejlepšího zástupce pro dva vybrané datasety (dataset 2 a dataset 3). V tomto případě se jedná o algoritmus strojového učení, jehož hyperparametry byly nastaveny podle výsledků vzešlých z evolučního algoritmu. Ve všech případech tento zástupce dosahuje výborných výsledků, kde metrika MFPR dosahuje velmi nízkých až nulových hodnot. Takto nastavený systém má poté v reálném provozu velmi nízké procento falešných poplachů.

## 6.5 Interpretace anomálií detekovaných pomocí algoritmů strojového učení

Tato podkapitola disertační práce je zaměřena na zhodnocení možnosti interpretace výsledků. V rámci předešlých experimentů byli identifikováni dva nejlepší zástupci pro detekci kybernetických útoků. Jedná se o neuronovou síť a algoritmus strojového učení IF. Avšak tyto dvě zvolené varianty se od sebe mírně odlišují z pohledu interpretace výsledků. V případě IF je nutné vytvořit druhý model algoritmu strojového učení Random forest (RF) k modelu IF. RF poté dostává výstup z IF, pokud byla identifikována anomálie. RF poté může identifikovat jednotlivé významnosti datových bodů. Všechny ostatní postupy jsou u obou algoritmů totožné.




Obr. 3: Interpretace výsledků neuronové sítě – dataset 2. [vlastní zdroj]

V Obr. 3 je zobrazen diagram postupu pro interpretaci výsledků v rámci neuronové sítě. Celý proces začíná získáním predikce výsledků pomocí modelu neuronové sítě. Následuje proces výpočtu hraniční hodnoty pro rozdělení anomálního chování od normálního provozu systému. V závislosti na hraniční hodnotě je možné identifikovat anomální hodnoty. Jestli není nalezena anomálie, poté provoz detekčního systému může pokračovat stejně i nadále. Pokud je však anomálie nalezena, pak je nutné ji identifikovat a interpretovat. Jak je v diagramu názorně ukázáno, definice hraničních hodnot se děje hned po predikci. Tento proces je vlastně součástí detekce anomálií. Lze také konstatovat, že v této fázi interpretace jsou již známé jednotlivé hodnoty pro dílčí datové body. Jedná se o rozdíl reálné hodnoty a predikované hodnoty pro každý bod. Při identifikaci anomálie následuje poté poměrně jednoduchý postup. Tedy identifikace anomálního datového bodu hlavní komponenty, která k němu přináleží. V prezentovaném případě se jedná o hlavní komponentu 2 (viz diagram). Poté reverzní metodou lze docílit výčtu atributů v rámci hlavní komponenty včetně jejich významností. Tedy jak významný je každý z dílčích atributu pro sestavení hlavní komponenty (PCA). Prostřednictvím tohoto postupu lze zprostředkovaně vypočítat významnost jednotlivých atributů.

Výsledky takto získané mohou být podrobeny detailnější analýze, ze které by následně měla být vytvořena nezbytná opatření pro úpravu chodu sledovaného systému, popřípadě přijaty úpravy detekčního systému v případě falešné identifikace. V případě pozitivní identifikace kybernetického útoku by měla být přijata taková opatření, která zabraňují v pokračování útoku nebo zmírňují jeho dopady. Tato disertační práce nebyla svým pojetím koncipována, tak aby tyto body vyřešila. V řadě případů je nutná hluboká znalost chráněného systému a jeho procesů pro efektivní analýzu a také přijetí opatření z toho vycházející. V těchto případech je poté nutná úzká koordinace s technickými pracovníky chráněného systému pro vyřešení této problematiky.

# 7. PŘÍNOSY PRÁCE PRO VĚDU A PRAXI

V rámci předložené disertační práce byl proveden výzkum v oblasti detekce anomálií pomocí algoritmů a metod strojového učení.

## 7.1 Přínos pro vědu

Hlavním cílem disertační práce bylo vytvoření systému pro detekci anomálií v oblasti ICS. Tento systém byl od počátku koncipován s ohledem na specifika ICS systémů, při zachování možnosti interpretace výsledků. Tyto charakteristiky spolu s určitou adaptabilností navrhovaného řešení lze považovat za jeden z hlavních přínosů pro vědeckou komunitu. K dosažení tohoto cíle bylo zapotřebí multidisciplinárního přístupu, ve kterém se spojují znalosti spadající do oblasti kybernetické bezpečnosti, umělé inteligence, detekce anomálií, dolování dat, úpravy datasetů, optimalizace, multikriteriálního hodnocení, ale i oblasti průmyslových řídicích systémů. V rámci těchto definovaných oblastí lze očekávat přínosy pro vědeckou komunitu.

Jako další přínos pro vědu je potvrzení aplikovatelnosti představeného systému detekce anomálií i pro reálné systémy. Tedy aplikovatelnost tohoto systému v reálném prostředí při využití reálných dat. Využití detekce anomálií na základě algoritmů strojového učení se ukázalo jako velmi slibná oblast, která povede ke zvýšení kybernetické bezpečnosti i velmi komplexních systémů.

Výsledky disertační práce naznačují vhodné kombinace optimalizačních algoritmů, algoritmů strojového učení a technik pro úpravu dat v oblasti detekce kybernetických útoků. Z výsledků lze také vyvodit nevhodnost využití některých algoritmů strojového učení pro řešení zadané problematiky. Jedná se zejména o algoritmus OCSVM, který je často využíván odbornou komunitou pro detekci kybernetických útoků v rámci ICS systémů.

Poslední sekce přínosu pro vědeckou komunitu je zaměřena na interpretaci výsledků systému pro detekci anomálií. Při postupu interpretace významnosti jednotlivých atributů je využito reverzních postupů, které se z části odvíjejí od technik využitých v první sekci (techniky pro úpravu dat). Navržený postup lze aplikovat na řadu oblastí strojového učení, kde je využito různých vstupních dat o rozdílné dimensionalitě.

## 7.2 Přínos pro praxi

Kybernetická bezpečnost kritické infrastruktury se stala jednou z hlavních bezpečnostních otázek současnosti. Navržený systém pro detekci anomálií v oblasti

ICS je vytvořen s ohledem na aplikaci v reálném prostředí. Jednotlivé zvolené postupy jsou tudíž zvoleny takovým způsobem, jenž umožnuje nasazení popisovaného řešení v rozličných ICS systémech. Tento postup umožňují optimalizační techniky. Ty jsou využity pro výběr vhodného nastavení (hyperparametrů) algoritmů strojového učení. Toto nastavení i díky využitému multikriteriálnímu hodnocení umožňuje určité adaptace pro každý dílčí systém ICS.

Většina řešení pro ochranu před kybernetickými útoky je řešena na bázi detekce pomocí signatur nebo metod, které vyžadují přesně definovanou množinu kybernetických útoků, proti kterým je poté dotyčný systém chráněn. Navrhované řešení popsané v disertační práci, však využívá informace získané v rámci normálního chodu sledovaného systému. Na základě tohoto chování je vytvářen prediktivní model, který je dále využit pro detekci kybernetických útoků. Tento rozdíl oproti klasickému pojetím detekce kybernetických útoků pomocí signatur umožňuje systému detekce anomálií zásadní výhody při detekci kybernetických útoků. Systém detekce anomálií nepotřebuje žádnou databázi (signatur) kybernetických útoků. Takový systém umožňuje minimalizovat náklady spojené s periodickými aktualizacemi databází (signatur) kybernetických útoků. Druhým významným přínosem tohoto představeného řešení je detekce dosud neznámých kybernetických útoků. V případě detekce pomocí signatur tato událost nastává ve dvou případech. Při pozdní aktualizaci databáze signatur nebo v případě kdy kybernetický útok vůbec nebyl identifikován odbornou komunitou (případ „Zero Day Attack“). Navržený systém detekce anomálií byl od počátku vytvářen, aby nebyl negativně ovlivněn těmito dvěma případy. Aplikace takto navrženého systému detekce anomálií v praxi nejenom zvýší efektivnost detekce kybernetických útoků, ale také sníží náklady na provoz kybernetické ochrany ICS systému.

Z důvodu ověření a validace systému detekce anomálií byl uskutečněn rozsáhlý výzkum. Byla provedena řada experimentů, která se především zaměřila na optimalizaci algoritmů strojového učení prostřednictvím pěti metrik $M_{F1}$, $M_{MCC}$, $M_{Prec}$, $M_{FPR}$, Čas. Přičemž právě metrika $M_{FPR}$ (vyjadřuje falešné detekce kybernetických útoků) je nejdůležitější metrikou pro praxi. Výsledná řešení vykazují velmi malé až nulové hodnoty metriky $M_{FPR}$. Tyto výsledky podporují možnost nasazení systému detekce anomálií v reálném provozu. Právě absence falešných klasifikací nezatěžuje chráněný systém falešnými poplachy, které by mohly ohrozit kontinuitu provozu celého systému ICS. Zbylé metriky $M_{F1}$, $M_{MCC}$, $M_{Prec}$, Čas zajištují pomocí multikriteriálního hodnocení TOPSIS vyváženost výsledného modelu.

# 8. ZÁVĚR

Disertační práce byla zaměřena na ochranu systémů ICS, které se díky digitalizaci a Průmyslu 4.0 stávají nezbytnou součástí moderní společnosti. Tento trend závislosti na systémech ICS bude v budoucnu posilovat. Lze očekávat implementace ICS v řadě sektorů kritické infrastruktury státu. Z dosavadního vývoje v oblasti kybernetické bezpečnosti vyplývá narůstající zájem, jak státních, tak privátních aktérů o systémy ICS. Tento nárůst je zapříčiněn významností těchto systémů pro moderní společnost. Z tohoto důvodu je kybernetická bezpečnost ICS systémů nezbytná pro zachování dostupnosti kritických služeb pro obyvatelstvo. V rámci disertační práce byl představen ucelený systém pro detekci anomálií založený na strojovém učení. Tento systém byl od počátku navrhován tak aby splňoval následující požadavky:

1. Detekování neznámých kybernetických útoků.
2. Škálovatelnost systému.
3. Možnost reálného využití v systémech ICS.
4. Možnost interpretace výsledků.

Tyto čtyři požadavky byly zásadní pro tvorbu prezentovaného systému. První bod byl splněn již výběrem vhodného typu algoritmů strojového učení. Využití kombinace učení s učitelem a učení bez učitele zajišťuje tvorbu klasifikačních modelů výhradně prostřednictvím dat normálního provozu sledovaného systému. Druhý bod byl zajištěn pomocí transformace dat pomocí algoritmu PCA. Tento postup zajišťuje zpracování detekcí anomálií pro teoreticky velmi rozsáhlé systémy ICS. Třetí bod byl z pohledu počtu experimentů a časového hlediska nejobtížnějším. Bylo potřeba provést velmi velký počet experimentů, které se vztahovaly k úpravě datasetů, optimalizaci algoritmů strojového učení a k samotnému výběru nejlepšího řešení podle jednotlivých metrik. Výsledky ukázaly, že využití popisovaného řešení je možné uplatnit v reálných systémech. Zvláště pak v případě algoritmu IF.

Závěrečná kapitola byla zaměřena na interpretaci detekovaných anomálií. Bylo využito reverzního postupu pro získaní nejvýznamnějších atributů pro každý klasifikovaný anomální záznam. Takto klasifikované výsledky byly poté předány pro detailnější analýzu, které jsou následně nezbytné pro tvorbu opatření pro úpravu chodu sledovaného systému. Tato disertační práce si však svým pojetím neklade za cíl finální interpretaci původu, typu kybernetického útoku, popřípadě detekci útočníka. V řadě případů je nutná hluboká znalost chráněného systému a jeho procesů pro efektivní analýzu a také přijetí opatření z toho vycházející.

# 9. SEZNAM POUŽITÉ LITERATURY


[1] FRANK, Alejandro Germán; DALENOGARE, Lucas Santos; AYALA, Néstor Fabián. Industry 4.0 technologies: Implementation patterns in manufacturing companies. International Journal of Production Economics, 2019, 210: 15-26.

[2] FALLIERE, Nicolas; MURCHU, Liam O.; CHIEN, Eric. W32. stuxnet dossier. White paper, Symantec Corp., Security Response, 2011, 5.6: 29.

[3] STOUFFER, Keith; FALCO, Joe; SCARFONE, Karen. Guide to industrial control systems (ICS) security. NIST special publication, 2011, 800.82: 16-16.

[4] DEWA, Zibusiso; MAGLARAS, Leandros A. Data mining and intrusion detection systems. International Journal of Advanced Computer Science and Applications, 2016, 7.1: 62-71.

[5] SOKOLOV, Alexander N.; PYATNITSKY, Ilya A.; ALABUGIN, Sergei K. Research of classical machine learning methods and deep learning models effectiveness in detecting anomalies of industrial control system. In: 2018 Global Smart Industry Conference (GloSIC). IEEE, 2018. p. 1-6.

[6] GOODFELLOW, Ian, et al. Deep learning. Cambridge: MIT press, 2016.

[7] LEMAY, Antoine; FERNANDEZ, José M. Providing {SCADA} network data sets for intrusion detection research. In: 9th Workshop on Cyber Security Experimentation and Test ({CSET} 16). 2016.

[8] GOH, Jonathan, et al. A dataset to support research in the design of secure water treatment systems. In: International Conference on Critical Information Infrastructures Security. Springer, Cham, 2016. p. 88-99.

[9] MORRIS, Thomas H.; THORNTON, Zach; TURNIPSEED, Ian. Industrial control system simulation and data logging for intrusion detection system research. 7th annual southeastern cyber security summit, 2015, 3-4.

[10] LIPTON, Zachary C.; BERKOWITZ, John; ELKAN, Charles. A critical review of recurrent neural networks for sequence learning. arXiv preprint arXiv:1506.00019, 2015.

[11] BERGSTRA, James S., et al. Algorithms for hyper-parameter optimization. In: Advances in neural information processing systems. 2011. p. 2546-2554.

[12] DEMŠAR, Janez. Statistical comparisons of classifiers over multiple data sets. Journal of Machine learning research, 2006, 7.Jan: 1-30.

# 10. SEZNAM OBRÁZKŮ



# 11. SEZNAM TABULEK



# 12. SEZNAM POUŽITÝCH ZKRATEK

| | |
|---|---|
| ANN | Artificial Neural Network |
| CF | Cílová Funkce |
| DT | Decision Tree |
| FPR | False Positive Rate |
| GA | Genetický algoritmus |
| ICS | Industrial control system |
| ICS-CERT | Industrial control system – Computer Emergency Response Team |
| ICT | Information and communication technologies |
| IF | Random Forest |
| IoT | Internet of Thinks |
| KI | Kritická infrastruktura |
| LSTM | Long Short Term Memory |
| MCC | Matthews Correlation Coefficient |
| MH | Multikriteriálního Hodnocení |
| MTU | Master Terminal Unit |
| OCSVM | One-class Support Vector Machines |
| PCA | Principal Component Analysis |
| Prec | Precision |
| RS | Grid search |
| RTU | Remote Terminal Unit |
| SCADA | Supervisory Control And Data Acquisition |
| TOPSIS | Technique for Order of Preference by Similarity to Ideal Solution |

| | |
|---|---|
| TPE | Tree-structured Parzen Estimator |

# 13. PUBLIKAČNÍ AKTIVITY AUTORA

**Mezinárodní publikace:**


1) VÁVRA, Jan; HROMADA, Martin. An evaluation of cyber threats to industrial control systems. In: International Conference on Military Technologies (ICMT) 2015. IEEE, 2015. p. 1-5. ISBN 978-80-7231-976-3.
2) VÁVRA, Jan; HROMADA, Martin; JAŠEK, Roman. Specification of the current state vulnerabilities related to industrial control systems. International Journal of Online and Biomedical Engineering (iJOE), 2015, 11.5: 64-68. ISSN 1868-1646.
3) VÁVRA, Jan; HROMADA, Martin. Comparison of the intrusion detection system rules in relation with the SCADA systems. In: Computer Science On-line Conference. Springer, Cham, 2016. p. 159-169. ISBN 978-3-319-33620-6.
4) VÁVRA, Jan; HROMADA, Martin. Possibilities of the Search Engine Shodan in Relation to SCADA. In: SECURWARE 2016, The Tenth International Conference on Emerging Security Information, Systems and Technologies, 2016. p. 130-135. IARIA. ISBN 978-1-61208-493-0.
5) VÁVRA, Jan; HROMADA, Martin. Determination of optimal cluster number in connection to SCADA. In: Computer Science On-line Conference. Springer, Cham, 2017. p. 136-147. ISBN 978-3-319-57141-6.
6) VÁVRA, Jan; HROMADA, Martin. Evaluation of anomaly detection based on classification in relation to SCADA. In: 2017 International Conference on Military Technologies (ICMT). IEEE, 2017. p. 330-334. ISBN 978-1-5386-1988-9.
7) VÁVRA, Jan; HROMADA, Martin. Anomaly detection system based on classifier fusion in ics environment. In: 2017 International Conference on Soft Computing, Intelligent System and Information Technology (ICSIIT). IEEE, 2017. p. 32-38. ISBN 978-1-4673-9899-2.
8) VÁVRA, Jan; HROMADA, Martin. Novelty Detection System Based on Multi-criteria Evaluation in Respect of Industrial Control System. In: Computer Science On-line Conference. Springer, Cham, 2018. p. 280-289. ISBN 978-331991191-5.
9) VAVRA, Jan; HROMADA, Martin. Comparative Study of Feature Selection Techniques Respecting Novelty Detection in the Industrial Control System

Environment. Annals of DAAAM and Proceedings of the International DAAAM Symposium, 2018, 29. p. 1084-1091. ISBN 978-3-902734-20-4.

10) VAVRA, Jan; HROMADA, Martin. Optimization of the Novelty Detection Model Based on LSTM Autoencoder for ICS Environment. In: Proceedings of the Computational Methods in Systems and Software. Springer, Cham, 2019. p. 306-319. ISBN 978-3-030-30328-0.
11) VAVRA, Jan; HROMADA, Martin. Evaluation of Data Preprocessing Techniques for Anomaly Detection Systems in Industrial Control System. Annals of DAAAM & Proceedings, 2019, 30. p. 738-745.

**Tuzemské publikace:**

1) VAVRA, Jan. Optimization of Crisis Management in Municipality via GIS. In: Trilobit [online]. Univerzita Tomáše Bati ve Zlíně, Fakulta aplikované informatiky, 2015. č. 1/2015, ISSN 1804-1795.
2) VAVRA, Jan. Ochrana ICT před škodlivým působením blesku. In: Trilobit [online]. Univerzita Tomáše Bati ve Zlíně, Fakulta aplikované informatiky, 2015. č. 1/2015, ISSN 1804-1795.
3) VAVRA, Jan; HROMADA, Martin. Specifikace Kybernetických Incidentů Vztahujících se k ICS. In: Bezpečnostní technologie, systémy a management 2015: Sborník příspěvků 5. mezinárodní konference. 1. Zlín: Univerzita Tomáše Bati ve Zlíně, Fakulta aplikované informatiky, 2015, s. 1-6. ISBN 978-80-7454-559-7.
4) VAVRA, Jan; HROMADA, Martin. Zhodnocení Detekčních Metodologií IDS ve Vztahu k ICS. In Sborník příspěvků z mezinárodní konference MLADÁ VĚDA 2016. Ostrava: Sdružení požárního a bezpečnostního inženýrství, z.s., 2016, s. 466-471. ISBN 978-80-7385-177-4.
5) VAVRA, Jan; HROMADA, Martin. Umělá inteligence jako nástroj ochrany kritické infrastruktury. In Sborník příspěvků z mezinárodní konference MLADÁ VĚDA 2019. Ostrava: Sdružení požárního a bezpečnostního inženýrství, z.s., 2019, s. 89- 99. ISBN 978-80-7385-222-1.
6) VAVRA, Jan; HROMADA, Martin. Metodika pro výběr metod určených pro kvantifikaci penalizačních faktorů v oblasti konvergované bezpečnosti. In: Trilobit [online]. Univerzita Tomáše Bati ve Zlíně, Fakulta aplikované informatiky, 2019. č. 3/2019, ISSN 1804-1795.

# 14. ODBORNÝ ŽIVOTOPIS AUTORA

**Ing. Jan Vávra**

**Osobní údaje**
**Datum narození:** 13. 10. 1987

**Adresa:** Úprkova 1809, 68603 Staré Město (Česká republika)

**E-mail:** jvavra@utb.cz

**Tel.:** +420722691886

---

**Vzdělání**

**Dosažené vzdělání:**
vysokoškolské II. stupě (Magisterské) – Ing.

**09/2014–do současnosti**
Ph.D. student ve studijním oboru Inženýrská informatika
Univerzita Tomáše Bati ve Zlíně, Zlín (Česká republika)

**17/06/2014–18/06/2014**
Certifikát z oblasti základů elektronického zabezpečení objektů JABLOTRON ALARMS a.s.,
Zlín (Česká republika)

**09/2009–06/2014**
Inženýrský titul v oboru Bezpečnostní technologie, systémy a management
Univerzita Tomáše Bati, Zlín (Česká republika)

**09/2004–05/2008**
Maturita v oboru Technické lyceum
Střední průmyslová škola, Uherské Hradiště (Česká republika)

---

**Přehled aktivit během studia**

**1. 1. 2019 - 31. 12. 2019**
Řešitel projektu IGA (IGA/FAI/2019/002) Detekce kybernetických útoků v prostředí průmyslových řídicích systémů prostřednictvím strojového učení.

**1. 6. 2018 - 31. 7. 2018**
Pracovní stáž v zahraničí Erasmus+ - Holandsko - University of Twente.

**1. 1. 2018 - 31. 12. 2018**
Řešitel projektu IGA (IGA/FAI/2018/003) Konceptuální návrh metodiky detekce anomálií vztahující se k průmyslovým řídicím systémům.

**1. 1. 2017 - 31. 12. 2017**
Řešitel projektu IGA (IGA/FAI/2017/003) Evaluace detekčních metodologií ve vztahu ke SCADA systémům.

**1. 6. 2016 - 31. 7. 2016**
Pracovní stáž v zahraničí Erasmus+ - Itálie - University of Cagliari.

**1. 1. 2016 - 31. 12. 2016**
Řešitel projektu (IGA IGA/FAI/2016/014) - Specifikace ICS kybernetické bezpečnosti se zaměřením na IDPS.

**od 2016 do 2019**
Spoluřešitel projektu ev. no. VI20172019054 - Analytický programový modul pro hodnocení odolnosti v reálném čase z hlediska konvergované bezpečnosti.

**1. 3. 2015 - 31. 12. 2015**
Řešitel projektu IGA (IGA/FAI/2015/042) - Analýza kybernetické bezpečnosti v organizaci se zaměřením na ICS.

**20. 9. 2015 - 20. 12. 2015**
Studijní zahraniční pobyt Erasmus+ - Řecko - University of Peloponnese.

**od 2015 do 2018**
Spoluřešitel projektu ev. no. VI20152019049 "RESILIENCE 2015: Dynamické hodnocení odolnosti souvztažných subsystémů kritické infrastruktury.

Ing. Jan Vávra, Ph.D.

# Návrh a ověření systému detekce anomálií založeného na strojovém učení v průmyslových řídicích systémech

Design and verification of anomaly detection system based on machine learning in industrial control systems

Teze disertační práce

Vydala Univerzita Tomáše Bati ve Zlíně,

nám. T. G. Masaryka 5555, 760 01 Zlín.

Náklad: vyšlo elektronicky

Sazba: Ing. Jan Vávra, Ph.D.

Pořadí vydání: první

Publikace neprošla jazykovou ani redakční úpravou.

Rok vydání 2020

ISBN 978-80-7454-976-2