

## Detekce anomálií v síťové komunikaci

- Cílem projektu je ověřit možnosti detekce anomálií v průmyslové komunikaci IEC 104 pomocí statistických vlastností provozu.
- Projekt se skládá z několika částí:
  1. Analýza komunikace a výběr statistických vlastností (features) paketů IEC 104, které reprezentují typické chování provozu IEC 104. Příkladem takových vlastností je například velikost paketů, délka mezipaketové mezery, počet přenesených bytů či paketů za určitý čas apod.
  2. Extrakce vybraných parametrů (vlastností) paketů ze zvoleného souboru ve formátu PCAP do strojově čitelného formátu, například CSV.
  3. Vytvoření modelu statistického rozložení zvolených parametrů pomocí metod strojového učení, například Box Plot, pravidlo tří sigma, metoda One-Class SVM, časové řady, One-Class Neural Networks apod. Pro implementaci můžete použít dostupné knihovny, například [Scikit-learn](https://scikit-learn.org/).
  4. Natrénování modelu pomocí 2/3 vybraného datasetu. Poslední třetinu dat použijte k validaci detekce. Zjistěte počet FP, FN a vypočítejte matici záměn (confusion matrix). Spočítejte metriky Accuracy a Precision pro vyhodnocení přesnosti detekce.
  5. Zhodnocení výsledků. Ověřte vylepšení detekce úpravou vstupních parametrů detekčního systému a nastavením prahové hodnoty.
  6. Vytvoření technické zprávy (5-10 stran). V dokumentu popište jednotlivé kroky řešení a dosažené výsledky.
  7. Odevzdání řešení jako archiv ZIP přes IS FIT. Soubor se jménem xlogin.zip bude obsahovat následující soubory:
    - Readme.txt - jméno, login, popis přiložených souborů, stav řešení
    - Technická zpráva (soubor xlogin.pdf)
    - Upravená data (např. ve formátu CSV), které jste použili k trénování a ověřování detekce
    - Zdrojové kódy skriptů či aplikace pro zpracování dat a vytvoření modelu.
- Doporučení postup pro řešení je popsán v prezentaci, která je uložena v IS FIT u předmětu PDS v části projekty. Zde také najdete odkazy na dostupné články a nástroje.
- Základní informace k projektu dostanete na demo cvičení v pátek 18.2.2022 po přednášce. Prezentace k řešení je k dispozici u termínu Projekt v IS FIT. Dále je možné využít forum k projektu.