

Hodnotenie rizík (Neupravená verzia)

Implementácia digitálneho dvojčata 5G siete spolu s modelom ML prináša rôzne výzvy, ktoré môžu ovplyvniť presnosť predpovedí, stabilitu a efektívnosť celého systému. Identifikácia týchto problémov a návrh stratégií na ich zmiernenie sú jednou z kľúčových úloh tejto práce.

Nedostatočná kvalita vstupných dát je jedným z najpravdepodobnejších rizík, ktoré môžu viesť k nesprávnym výsledkom modelu. Chyby v dátach alebo ich nereprezentatívnosť, napríklad pri modeloch sieťovej prevádzky (traffic patterns), môžu narušiť presnosť predpovedí \cite{ML_traffic}. Formou zmiernenia je testovanie na rôznych dátových scenároch a aplikácia metód ako krížová validácia a ladenie hyperparametrov, ktoré minimalizujú riziko chýb spojených s podtrénovaním (underfitting) a pretrénovaním (overfitting) \cite{Nguyen}.

Zber kvalitných dát môže byť problematický, pretože mnohé scenáre je potrebné zachytiť v laboratóriu. Bez dostatočných dát môže model generovať neadekvátne predpovede, ktoré nebudú reprezentovať skutočnosť. Riešením môže byť generovanie syntetických údajov \cite{data_generating} a využitie dostupných datasetov z iných projektov \cite{datasets_telecom}, ktoré môžu čiastočne nahradiť reálne dáta a napomôcť k presnejším predpovediam.

Nezvyčajné scenáre, ako vysoké zaťaženie siete (peak loads), neštandardné správanie používateľov či poveternostné podmienky môžu narušiť schopnosť modelu adaptovať sa \cite{challenges_human_factor}. Ak tieto scenáre neobsahujú tréningové dáta, model nemusí byť pripravený na takéto situácie. Vzhľadom na čas, ktorý máme na získanie a spracovanie dát, tento problém nemusí byť vo finálnej implementácii vyriešený dostatočne, a preto jeho vyriešenie vyžaduje ďalšiu prácu a zber dát.

Kompatibilita systémov ako srsRAN, Open5GS, UERANSIM a nástrojov Docker nie je vždy automatizovaná, čo môže ovplyvniť celkovú integráciu práce \cite{challenges_human_factor}. Z tohoto dôvodu je vhodné začínať implementáciu na malých izolovaných komponentoch, za neustáleho testovania. Takéto testovanie, a celkový vývoj DT sú časovo veľmi náročné, čo potvrdili aj autori v \cite{USAirForce}. Tento fakt môže negatívne vplyvať na výsledok celej práce, nakoľko čas je jedným z kľúčových faktorov, ktoré majú vplyv na množstvo a kvalitu pozberaných dát určených na tréningovanie ML modelu.

Konfigurácia siete môže odhaliť citlivé informácie o 5G infraštruktúre či porušenie GDPR \cite{big-data-problems} pri nechcenom zachytení údajov o používateľoch \cite{challenges-technol}. Únik takýchto dát by ohrozil nielen bezpečnosť projektu, ale aj reálnej siete \cite{Dt_lot_data_worry_about}. Používanie .env súborov na uchovávanie citlivých premenných a simulovanie siete s fiktívnymi údajmi výrazne znižuje riziko úniku.

Experimentálna reprodukovateľnosť a integrácia (neupravená verzia)

Pre zabezpečenie experimentálnej reprodukovateľnosti a jednoduchšej integrácie bola zvolená kontajnerizovaná architektúra pomocou Docker a Docker Compose. Namiesto vytvárania vlastných konfiguračných súborov a obrazov bol použitý existujúci open-source repozitár \texttt{herlesupreeth/docker_open5gs} \cite{herlesupreeth}, ktorý poskytuje dockerizované prostredie pre komponenty siete 5G vrátane \texttt{Open5GS}, \texttt{UERANSIM}, ako aj voliteľnej integrácie s \texttt{Prometheus} a \texttt{Grafana} na zber a vizualizáciu metrík.

Repozitár bol naklonovaný z GitHubu a jednotlivé obrazy boli zostavené pomocou príkazov:

Niektoré súbory z originálnej implementácie boli upravené pre potreby tohoto projektu. Do yaml súboru, ktorým sa spúšťali niektoré komponenty boli pridané ďalšie služby tak, aby stačil jeden príkaz pre spustenie všetkých komponentov.

Do súboru `\textbf{deploy-all.yaml}` bola pridaná služba pre integráciu služby NetData `\cite{netdata}`. Najskôr bolo treba sa na službe zaregistrovať a vytvoriť pracovný priestor. NetData následne ponúka možnosť použiť ich API priamo pomocou Dockeru.

Spustenie kontajnerov bolo upravené tak, aby sa dalo spúšťať všetko jedným príkazom:

```
\begin{lstlisting}[language=bash,caption={Build pre jednotlivé docker images}, style=bashstyle] docker compose -f deploy-all.yaml up --build -d \end{lstlisting}
```

Táto voľba bola motivovaná cieľom maximalizovať znovupoužiteľnosť a minimalizovať konfiguračné chyby. Repozitár je aktívne udržiavaný a poskytuje konzistentné a overené nastavenia, ktoré výrazne urýchľujú vývojový proces. Hlavnou výhodou tohto prístupu je zníženie bariéry pre reprodukovateľnosť – akýkoľvek výskumník s podporovaným operačným systémom a Dockerom môže systém replikovať lokálne v priebehu niekoľkých minút. Nevýhodou je menšia kontrola nad interným nastavením kontajnerov, čo môže byť limitujúce pri pokročilejších úpravách.

\par{ Pre účely monitorovania spotreby systémových zdrojov bola do architektúry doplnená open-source platforma `\textbf{NetData}`. Táto služba bola nainštalovaná priamo do Docker hosta a nakonfigurovaná na sledovanie jednotlivých kontajnerov cez dostupné Docker plug-iny. NetData poskytuje prehľad o využití `\textbf{RAM}`, `\textbf{CPU}`, `\textbf{siete}`, ako aj o `\textbf{sietovej komunikácii jednotlivých kontajnerov}`, čo je kľúčové pre ladenie a optimalizáciu simulácií. }

Monitorovanie pomocou NetData zvyšuje transparentnosť experimentálneho prostredia a umožňuje identifikovať úzke miesta pri spustení viacerých UEs alebo vysokom trafficu. Zároveň je vďaka tomu možné porovnávať výkonnostné profily experimentov medzi rôznymi behmi, čo je kľúčovým predpokladom pre experimentálnu reprodukovateľnosť.

Udržateľnosť a environmentálny dopad (neupravená verzia)

Implementácia opatrení na zabezpečenie udržateľnosti projektu a minimalizácia jeho environmentálneho dopadu sú kľúčové pre zaistenie dlhodobej hodnoty, spoločenského prínosu a efektivity vypracovania tejto práce. Práca je navrhnutá s dôrazom na efektívne využívanie zdrojov, udržateľný softvérový a hardvérový návrh a životný cyklus 5G siete, ktorý minimalizuje potrebu používania fyzických zdrojov.

Jednou z hlavných prínosov DT je možnosť predikcie a optimalizácie, pričom ak je DT zostrojené správne, môže dopomôcť k redukcii spotreby elektrickej energie. `\cite{DT_edge_networks_IoT}`. Predikcia budúceho stavu siete umožňuje taktiež lepšiu správu záťaže (traffic load) a preťaženia (congestion), čo napomáha k znižovaniu nadmernej spotreby energie `\cite{malaysia_enviro}`. Navyše, DT eliminuje potrebu testovania na fyzických zariadeniach, čím sa minimalizuje spotreba rôznych materiálov `\cite{enviro_raw_materials}` a času potrebného na fyzické experimenty. Tento prístup je obzvlášť užitočný v prípade nasadzovania a testovania 5G technológií v oblastiach s nerozvinutou infraštruktúrou a obmedzenými výrobnými zdrojmi `\cite{huawei_i_cities}`.

Vývoj softvéru bol orientovaný na maximálnu efektivitu, čo zahŕňa optimalizáciu kódu na zníženie spotreby energie počas behu aplikácie a nasadenie projektu v prostredí Docker, čo umožňuje rýchlejšiu konfiguráciu a škálovateľnosť zariadení. Tieto opatrenia nielen znižujú environmentálny dopad `\cite{docker_enviro}\cite{docker_enviro_2}`, ale aj zvyšujú celkovú udržateľnosť projektu.

Modulárny dizajn `\cite{modular_sw}` projektu zabezpečuje, že aktualizácie a údržba nemajú vplyv na celkovú funkčnosť systému. Tento prístup znižuje potrebu kompletného prekonfigurovania alebo fyzických zásahov do chodu programu, čo prispieva k dlhodobej udržateľnosti. Takýto dizajn môže mať pozitívny vplyv na životné prostredie `\cite{modular_sw}` (Green Design).

Ako je vyššie uvedené, použitie prediktívnych modelov v DT môže viesť k zásadným environmentálnym prínosom \cite{enviro}. Okrem zníženia zaťaženie fyzickej infraštruktúry a menej častých aktualizácie fyzického hardvéru, môže mať za následok aj nižšiu spotrebu zdrojov a menšiu produkciu odpadu \cite{enviro_raw_materials}. Predikčné modely teda umožňujú efektívnejšie rozhodovanie s pozitívnym dopadom na životné prostredie.

Zamestnateľnosť (neupravená verzia)

Táto bakalárska práca, zameraná na rozvoj teoretických znalostí v oblasti digitálnych dvojčiat v spojení s praktickou implementáciou a optimalizáciou 5G sietí, vedie k rozšíreniu zručností vo viacerých kľúčových oblastiach technologického sektora. V neposlednom rade strojové učenie, použité na predikciu správania implementovaného digitálneho dvojčaťa, zasahuje aj do oblasti dátovej vedy.

Vďaka formátu práce autori prejdú celým cyklom realizácie projektu, od prieskumu technológií cez návrh až po implementáciu a testovanie. Týmto získajú ucelený a komplexný pohľad na vývoj a riadenie softvérových projektov, ako aj na plánovanie, organizáciu a efektívnu komunikáciu.

Takáto kombinácia technických, projektových a komunikačných schopností môže významne zvýšiť hodnotu autorov na trhu práce, najmä v budúcnosti, keďže problematika digitálnych dvojčiat a 5G sietí je stále viac žiadaná a nachádza uplatnenie v rôznych odvetviach.