



Hodnocení vedoucího závěrečné práce

Vedoucí práce:	Ing. Pavel Prochazka, Ph.D.
Student:	Anton Bushuiev
Název práce:	Neuronové site na grafech zohlednující časový aspekt v kybernetické bezpečnosti
Obor / specializace:	Znalostní inženýrství
Vytvořeno dne:	4. června 2021

Hodnotící kritéria

1. Splnění zadání

- ▶ [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Náročnost zadání je dána zejména dvěma aspekty. Za prvé, práce využívající času v grafových neuronových sítích pocházejí především z posledních několika málo let, tudíž udělat ucelenou rešerši vyžaduje přečtení a pochopení většího množství článků. Druhý aspekt komplikující praktickou část práce je aplikace metody v reálném světě se všemi důsledky typu málo anotovaných dat, nevyváženost tříd, velikost dat atd. Tyto komplikace bylo nutné reflektovat při návrhu experimentů pro ověření metod, jejich evaluaci a interpretaci.

Způsob provedení rešerše literatury odpovídá (splňuje) zadání první části práce. Systematickým popisem relevantní literatury student prokázal pochopení principů grafových neuronových sítí a rozsah rešerše pokrývá základní současné metody grafových neuronových sítí.

Praktická část práce se věnuje aplikaci metod na Cisco-Cognitive datasetu. V této části je modelována bezpečnost domén jako binární klasifikátor uzlu, která je vyhodnocována na information retrieval úloze. V experimentech jsou poté porovnány různé metody na dané úloze. Uvedené experimenty neprokazují žádnou výhodu využití informace o čase na prozkoumaném datasetu. Nicméně tento poznatek považuji za adekvátní výsledek zadání bakalářské práce - i když nelze přímo využít v současném stavu.

Jako drobný nedostatek považuji vynechání některých relevantních metod (např. Weg2weg přinášející velmi flexibilní reprezentaci). Nicméně jsem si vědom, že uvažování dalších metod by zvětšilo rozsah práce.

Nad rámec zadání práce obsahuje vlastní přístup uvedený v sekci 3, kde je navržena konkrétní agregace v grafové neurovnové síti využívající čas. I když jsme později našli popis tohoto přístupu v literatuře, oceňuji studentovu schopnost přijít s návrhem a implementací vlastní metody, která logicky rozšiřuje uvedenou literaturu.

2. Písemná část práce

85 /100 (B)

K prvním dvěma sekcím po formální a obsahové stránce nemám vážnější připomínky. Práce je logicky strukturovaná, systematicky prezentuje současné způsoby použití grafových neuronových sítí s referencemi na relevantní literaturu. V praktické části by mohlo být lépe vysvětleno proč je bezpečnost domén vyhodnocována pomocí information retrieval přístupu. Dále by práce mohla obsahovat více referenčních datasetů s ukázkou, jak konkrétně využívají časovou informaci referenční metody.

3. Nepísemná část, přílohy

95 /100 (A)

Implementace metod je napsána v Pythonu v Jupyter noteboocích. Vzhledem k tomu, že se jedná hlavně o rychlé prototypování, považuji tento výstup za zcela odpovídající.

4. Hodnocení výsledků, jejich využitelnost

85 /100 (B)

Práce obsahuje 3 výstupy

- 1) Systematický přehled grafových neuronových sítí využívajících čas (Sekce 1, 2).
- 2) Vlastní nový poznatek - Time Series GNN (Sekce 3)
- 3) Implementace a porovnání metod na problému určování bezpečnostního rizika domén

Vzhledem k povaze zadání není ambice přímo zlepšit produkt. V dostupné literatuře se téměř nevyskytují grafové neuronové sítě využívající čas pro problémy internetové bezpečnosti. Považuji tedy výstupy práce spíše jako vstupní informace/poznatky pro další výzkum tímto směrem. Akceptovatelným výsledkem může být také poznatek, že informace o času opravdu nepřináší žádnou hodnotu pro danou úlohu. V tomto smyslu vynívají výsledky a následná diskuze. Přestože byly experimenty navrženy s konzultací, mohl být věnován větší prostor diskuzi jak konkrétní experimenty odpovídají na zadání práce a jaké případné další experimenty/metody by mohly být navrženy. Dále by práce mohla zmínit limitaci provedené evaluace. Závěr, který si z práce odnášíme, tedy není plně korespondující se závěrem popsáním v bakalářské práci, což může zkreslovat její skutečný přínos.

5. Aktivita studenta

- [1] výborná aktivita
- [2] velmi dobrá aktivita
- [3] průměrná aktivita
- [4] slabší, ale ještě dostatečná aktivita
- [5] nedostatečná aktivita

Ke spolupráci nemám výhrady. Student průběžně plnil zadání a přicházel s vlastními nápady.

6. Samostatnost studenta

- [1] výborná samostatnost
- [2] **velmi dobrá samostatnost**
- [3] průměrná samostatnost
- [4] slabší, ale ještě dostatečná samostatnost
- [5] nedostatečná samostatnost

Ze začátku bylo nutné seznámit studenta s problémem a dostupnými daty. Během semestru jsme se snažili dát hromady návrh experimentů a způsob jejich evaluace, který by mohl demonstrovat výhodu využití informace o čase. Po návrhu experimentů a způsobu evaluace jsme již diskutovali jenom jejich výsledky, na kterých student samostatně pracoval.

Celkové hodnocení

95 /100 (A)

I přes drobné výhrady k závěrům výrazně převažují pozitiva s ohledem na rozsah a kvalitu odevzané bakalářské práce.

Instrukce

Splnění zadání

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Nepísemná část, přílohy

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

Hodnocení výsledků, jejich využitelnost

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Aktivita studenta

V souvislosti s průběhem a výsledkem práce posudte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven.

Samostatnost studenta

V souvislosti s průběhem a výsledkem práce posudte schopnost studenta samostatně tvůrčí práce.

Celkové hodnocení

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.