

Záměr týmového projektu

Studijní program: Informatika – Softwarové a datové inženýrství

Typ projektu: Výzkumný projekt (NPRG070)

Student-řešitel: Oskar Razyapov

Supervizor: Martin Nečaský

Konzultant:

Název a téma projektu: Optimalizace zobrazování větších grafů

Návrh technického řešení, harmonogram, milníky a ostatní údaje:

Motivace:

Výzkumný tým Nečaský, Klímek, Škoda, Stenclák pracuje na výzkumném projektu, který se jmenuje „Prohlížeč znalostních grafů“.

Znalostní graf je datová struktura ve formě orientovaného matematického grafu. To znamená, že je tvořen uzly a orientovanými hranami. Každý uzel ve znalostním grafu reprezentuje určitou entitu reálného světa. Vztahy mezi jednotlivými entitami jsou pak reprezentovány jako orientované hrany. Kromě toho je každá hrana označena názvem reprezentovaného vztahu.

Prohlížeč znalostních grafů umožňuje běžným uživatelům prozkoumávat znalostní grafy pomocí předem nakonfigurovaných, jednoduchých, interaktivních pohledů.

Bohužel, často znalostní grafy jsou dost velké, anebo obsahují příliš mnoho detailů, a proto jsou obtížné na vizualizaci a neumožňují snadný vizuální průzkum a pochopení běžným uživatelům.

Popis projektu:

V rámci tohoto projektu student navrhne, naprogramuje a otestuje rozšíření stávajícího prohlížeče znalostních grafů.

Zaměří se na implementaci technik, které zjednoduší velký graf tak, aby za prvé jeho vizualizace a spouštění na něm layoutovacích algoritmů bylo rychlé a jednoduché, a za druhé, aby byl pro běžné uživatele dobře přehledný a snadno pochopitelný.

Jednou z možných technik zjednodušení grafu je kondenzace velkého původního grafu do malého, syntetického a vysoce informativního grafu. Při kondenzaci, z každé silně souvislé komponenty grafu se vytvoří jeden uzel, reprezentující danou komponentu. Výhodou této techniky je nejen snadnější ukládání, vizualizace a načítání grafů, ale také zrychlení algoritmů grafů.

Dalšími technikami jsou sparsifikace a zhrubnutí (coarsening). Sparsifikace snižuje počet hran při aproximaci párových vzdáleností. Bohužel, se stává mnohem méně slibnou při zjednodušování grafů, když jsou uzly také spojeny s atributy, protože sparsifikace nesnižuje atributy uzlů. Zhrubnutí snižuje

počet uzlů s podobnými omezeními typicky seskupováním původních uzlů do super-uzlů a definováním jejich spojení.

Taky lze použít metody grafové sumarizace, které uživatelům pomáhají extrahovat a pochopit základní informace v grafech. Metoda by měla uživatelům umožnit svobodný výběr atributů a vztahů, které je zajímají, a poté využít těchto funkcí k vytvoření malého a informativního souhrnného grafu, který odhaluje základní charakteristiky uzlů a jejich vztahů v původním grafu.

Výstup projektu:

V případě úspěšného řešení projektu bude výstupem rozšíření existujícího projektu o podporu vizualizaci a prozkoumání velkých znalostních grafů společně s textem pro publikaci článku, popisujícím využití techniky a příklady jejich použití.

Platforma, technologie:

- TypeScript, React.js
- SPARQL, RDF, JSON

Milníky a harmonogram:

Název milníku	Harmonogram
Detailní návrh možných řešení a ukázka toho, jak by mohl vypadat graf po zjednodušení	M01-M02
Implementace možných rozšíření a jejich integrace do existujícího prohlížeče znalostních grafů	M03-M05
Testování rozšíření	M06
Sepsání dokumentace a publikace	M07-M08
Odevzdání	M09