



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Název projektu: Mezinárodní centrum pro informaci a neurčitost

Registrační číslo: CZ.1.07/2.3.00/20.0060

Zpráva z účasti na stáži

Datum konání stáže: 2. 5. 2014 – 25.5.2014
Navštívené pracoviště: UTEP, El Paso, USA
Zahraniční garant: prof. Vladik Kreinovich
Účastník stáže: Mgr. Lukáš Havrlant

Stručný popis navštíveného pracoviště

University of Texas at El Paso byla založena v roce 1914, v současnosti ji navštěvuje 23 003 studentů, z toho 3 532 doktorandů a postdoců. Součástí univerzity je fakulta College of Engineering, která dále obsahuje 8 kateder. Jednou z nich je katedra Computer Science, na které jsem navštívil profesora Vladika Kreinoviche. Na katedře působí několik různých výzkumných skupin, z nich některé jsou propojené s armádou, i díky tomu, že v El Pasu se nachází jedna z největších vojenských základen Fort Bliss. Nejspíše nejvýznamější skupina, *Army High Performance Computing Research Center (AHPCRC)*, se například zabývá simulací a modelováním vícejádrových procesorů a hybridní optimalizací parametrických heuristických problémů.

Hlavní částí výzkumu Profesora Vladika Kreinoviche pak je intervalová logika a hledání obecných metod na odhadnutí přesnosti výsledků datových algoritmů. Dále se profesor Kreinovich zabývá matematickými základy automatického ovládání s cílem najít způsob jak převést znalosti experta do znalostní báze dané aplikace.

Průběh stáže

V průběhu stáže jsme s profesorem Kreinovichem a jeho kolegy zejména studovali teoretické základy tf-idf algoritmu, což je algoritmus, který slouží k řazení dokumentů nebo například k hledání klíčových slov dokumentů. Algoritmus funguje tak, že spočítá počet výskytů daného slova v dokumentu a vynásobí tuto hodnotu s logaritmem podílu počtu všech slov v dokumentu a počtem všech dokumentů, které dané slovo obsahují. Kromě této základní verze ještě existuje mnoho variací tohoto algoritmu, nicméně všechny variace jsou velmi podobné a

ani jedna neposkytuje žádné nové teoretické vysvětlení toho, proč by měl algoritmus fungovat.

Současná nabízená teoretická zdůvodnění funkčnosti algoritmu jsou buď nepřesvědčivá nebo příliš zdlouhavá. S profesorem jsme tak zkoumali alternativní teoretické vysvětlení algoritmu. Po prozkoumání stávající literatury jsme přišli s novým a jednodušším vysvětlením, založeném na statistice, využívající centrální limitní větu.

V rámci stáže jsem si dále prohlédl celou univerzitu, seznámil se s většinou lidí působících na katedře, přednesl jsem jen přednášku a společně s profesorem jsme napsali jeden vědecký článek.

Přednášky na navštíveném pracovišti

L. Havrlant: *Search Engine Based on Formal Concept Analysis*

Tuto přednášku jsem přednesl na neformálním vědeckém semináři v průběhu mé stáže. Přednášel jsem o své diplomové práci — webový vyhledávač založený na FCA. Přednášky se zúčastnilo i několik studentů nižších ročníků, takže samotná přednáška byla více praktického charakteru s důrazem na ukázkou funkčnosti vyhledávače a jednoduchého popisu teorie, která stojí za vyhledávačem. Během následující diskuse jsem se pak dověděl, že podobný způsob napovídání dotazů používá i společnost Google ve svém stejnojmenném vyhledávači.

Publikace rozpracované během stáže

V návaznosti na uskutečněné vědecké diskuse jsme s profesorem Kreinovichem napsali vědecký článek popisující náš nový teoretický model, který stojí za algoritmem tf-idf. Článek detailně rozebírá náš nový postup založený na jednoduché statistice a centrální limitní větě a srovnává nově nabízený pohled se současnými pohledy. Článek by se měl, pokud bude schválen, objevit v časopise *International Journal of General Systems*.

Navázání kontaktů

Navázal jsem kontakt s profesorem Vladikem Kreinovichem, seznámil jsem se s jeho výzkumem a sám jsem ho seznámil se svým výzkumem. Společně jsme pak napsali jeden článek a není vyloučena další spolupráce nad podobnou tematikou.

Shrnutí stáže

Stáž plně splnila svůj účel. V rámci stáže jsem získal řadu nových poznatků a doplnil si znalosti o aktuálních trendech v dané oblasti výzkumu. Tyto poznatky budou dále předány cílové skupině formou odborných seminářů.

Fotografická dokumentace



Fotografie z neformálního vědeckého semináře uskutečněného v průběhu stáže. Na snímku Mgr. Havrlant (vlevo) a prof. Kreinovich (vpravo).

Havrlant