

Posudek oponenta habilitační práce

Masarykova univerzita

Fakulta Fakulta informatiky MU

Habilitační obor Informatika

Uchazeč Mgr. Jan Outrata, Ph.D.

Pracoviště Univerzita Palackého v Olomouci

Habilitační práce Computing and Applying Formal Concepts: Algorithms and Methods

Oponent Prof. RNDr. Jan Rauch, CSc.

Pracoviště Fakulta informatiky a statistiky, Vysoká škola ekonomická v Praze, Praha

Předložená habilitační práce shrnuje a komentuje výsledky uchazeče a jeho spolupracovníků dosažené při výzkumu a aplikacích formální konceptuální analýzy. Práce se skládá z teoretické části a ze souboru deseti článků. Teoretická část má tři kapitoly. První obsahuje stručný přehled formální konceptuální analýzy včetně definic základních pojmů a základní věty o konceptuálních svazech. Druhá a třetí kapitola jsou věnovány výsledkům výzkumu ve dvou oblastech týkajících se formální konceptuální analýzy. Jedná se o algoritmy pro výpočet formálních konceptů a o aplikace formální konceptuální analýzy (FCA).

Součástí první kapitoly je i seznam článků, které tvoří druhou část předložené práce. U každého článku je mimo jiné uvedeno, ke kterým odstavcům ve které kapitole práce se článek vztahuje. U třech článků je uchazeč jediným autorem. U jednoho článku jsou dva spoluautoři, podíl uchazeče je u každého článku 50%. U čtyřech článků jsou tři spoluautoři, podíl uchazeče je dvakrát 33% a dvakrát 40%. U dvou článků jsou čtyři spoluautoři, podíl uchazeče je jednou 65% a jednou 75%. Čtyři z článků jsou publikovány v impaktovaných časopisech a šest článků ve sbornících mezinárodních konferencích. U jednotlivých článků je také uveden počet citací bez autocitací. Celkem se jedná o 19 citací ve WoS a 41 citací ve Scopusu, krom toho jsou i další citace.

Druhá kapitola se týká algoritmů pro výpočet formálních konceptů. Kapitola začíná podrobným přehledem současného stavu vývoje těchto algoritmů. Dále jsou představeny nové algoritmy vyvinuté na pracovišti uchazeče. Jedná se o tři algoritmy souhrnně nazývané *CbO-family* algoritmy a dále o algoritmus *attribute sorting* inspirovaný výsledky souvisícími s optimalizací CbO-algoritmů. Poté jsou uvedeny výsledky testů nových algoritmů na čtyřech často používaných datových sadách. Výsledky ukazují významné zlepšení efektivity nových algoritmů v porovnání se známými algoritmy pro výpočet formálních konceptů. V závěru kapitoly jsou uvedeny náměty na další výzkum. Ke druhé kapitole se vztahují články

- Krajca P., Outrata J., Vychodil V.: Parallel Recursive Algorithm for FCA. In: Belohlavek R., Kuznetsov S. O. (Eds.): CLA 2008: Proceedings of the Sixth Int. Conf. on Concept Lattices and Their Applications, 71- 82, 10/2008. CEUR WS, Vol. 433,
- Krajca P., Outrata J., Vychodil V.: Parallel Algorithm for Computing Fixpoints of Galois Connections. Annals of Mathematics and Artificial Intelligence 59(2)(2010), 257- 272.
- Krajca P., Outrata J., Vychodil V.: Advances in algorithms based on CbO. In: Kryszkiewicz M., Obiedkov S. (Eds.): CLA 2010: Proceedings of the 7th Int. Conf. on Concept Lattices and Their Applications, 325-337, CEUR WS, Vol. 672

- Outrata J., Vychodil V.: Fast Algorithm for Computing Fixpoints of Galois Connections Induced by Object-Attribute Relational Data. *Information Sciences* 185(1)(2012), 114-127.
- Krajca P., Outrata J., Vychodil V.: Computing formal concepts by attribute sorting. *Fundamenta Informaticae* 115(4)(2012)

Třetí kapitola popisuje dvě aplikace FCA. První aplikace se týká rozhodovacích stromů. Je navržena a experimentálně testována metoda tvorby rozhodovacích stromů pomocí FCA. Experimenty ukazují velmi dobré výsledky ohledně přesnosti vytvořeného stromu ve srovnání se známými algoritmy (ID3, C4.5 a další). Je také navržena řada námětů pro další výzkum v oblasti konstrukce rozhodovacích stromů pomocí FCA. Druhá aplikace se týká předzpracování dat pro data mining. Je navržena a testována metoda pro řešení úlohy *feature extraction* pomocí FCA. Metoda využívá faktorizaci booleovských matic. Je porovnávána kvalita rozhodovacích stromů vytvořených pomocí původních atributů a stromů vytvořených pomocí atributů vybraných pomocí navržené metody. Pro indukci stromů jsou použity algoritmy ID3 a C4.5. Ukazuje se, že stromy vytvořené pomocí atributů vybraných FCA poskytují lepší výsledky. I pro tuto metodu jsou uvedeny náměty na další výzkum. Ke kapitole se vztahují články

- Bělohlávek R., De Baets B., Outrata J., Vychodil V.: Inducing decision trees via concept lattices. In: Diatta J., Eklund P., Liquire M. (Eds.): *Proc. CLA 2007*, 274-285, Montpellier, France, 10/2007. CEUR WS, Vol. 331
- Outrata J.: Inducing decision trees via concept lattices. In: Trappl R. (Ed.): *Cybernetics and Systems 2008: Proceedings of the 19th European Meeting on Cybernetics and Systems Research*, 9-14
- Bělohlávek R., De Baets B., Outrata J., Vychodil V.: Inducing decision trees via concept lattices. *Int. Journal of General Systems* 38(4)(2009), 455-467.
- Outrata J.: Preprocessing input data for machine learning by FCA. In: Kryszkiewicz M., Obiedkov S. (Eds.): *CLA 2010: Proceedings of the 7th Int. Conf. on Concept Lattices and Their Applications*, 187-198, Sevilla, Spain, 10/2010. CEUR WS, Vol. 672,
- Outrata J.: Boolean factor analysis for data preprocessing in machine learning. In: Draghici S., Khoshgoftaar T. M., Palade V., Pedrycz V., Wani M. A., Zhu X. (Eds.): *Proceedings of The Ninth Int. Conf. on Machine Learning and Applications (ICMLA 2010)*, 899-902, Washington, D.C., USA, 12/2010.

Lze shrnout, že habilitační práce představuje významné výsledky výzkumu v důležité oblasti informatiky – formální konceptuální analýze dosažené na pracovišti uchazeče. Jejich význam je doložen mimo jiné 59 citacemi ve WoS a Scopusu. Z předložených podkladů vyplývá, že uchazeč se významným způsobem podílel na těchto výsledcích.

Dotaz oponenta:

Naznačte možnosti použití FCA při dolování asociačních pravidel z dat.

Závěr

Habilitační práce Jana Outraty „Computing and Applying Formal Concepts: Algorithms and Methods“ **splňuje** požadavky standardně kladené na habilitační práce v oboru Informatika.

Praha, dne 2. 6. 2016