

Provozně ekonomická fakulta Mendelovy univerzity v Brně

Ústav informatiky

Doc. Ing. Ivana Rábová, Ph. D.

Oponentský posudek na disertační práci Ing. Eva Zámečníková

Formal model of decision making process for high-frequency data processing

Úvod

Předložená disertační práce je zaměřena na oblast využití informatiky pro stanovení optimálních obchodních strategií, zabývá se problematikou vysokofrekvenčních časových řad a přispívá k rozvoji moderní a progresivní technologie Complex Event Processing (CEP).

Myšlenka podpory predikce vysokofrekvenčních časových řad (zde kategorizována jako Big Data) a výběr platformy pro jejich zpracování tak, aby odpovídaly podmínkám pro robustnost, škálovatelnost a zpracování požadavků v reálném čase je využitelná v praxi pro předpovídání, odhad trendů, případně vyhledání nových vzorů v evoluci dat. Obecně lze říci, že analýza časových řad, včetně predikce jejich budoucího chování, se postupně stává součástí běžných podnikových procesů ve finančních ale i ostatních institucích.

Z těchto důvodů považuji téma disertační práce za inspirující a využitelné, z pohledu vědeckého přínosu (pokud je mi známo) jde o novou metodu, nový přístup, nové pojetí dané problematiky s otevřenou budoucností.

Cíl práce

Autorka si vytyčila za svůj hlavní cíl rozšířit oblast zpracování a řízení vysokofrekvenčních časových řad využitím platformy komplexního řízení událostí. Svého cíle autor dosáhne naplněním dvou dílčích cílů, které jsou srozumitelně vyjádřeny. Navržená metoda pro formalizaci podnikových pravidel jako základ pro znalostní bázi nové a rozšiřující komponenty v CEP nástroji ESPER bude aplikována a otestována, bude provedeno hodnocení navrženého modelu v praxi a výsledky měření interpretovány.

Vzhledem k rostoucímu významu této disciplíny, komplexního zpracování událostí, považuji tento cíl za aktuální, formou a přístupem ke zpracování za disertabilní a adekvátní z pohledu zaměření v oboru Výpočetní technika a informatika.

Obsah a struktura práce

Práce obsahuje teoretickou a aplikační část, kapitoly jsou vhodně tematicky členěné a logicky navazující, téma je rozsáhlé, košaté a komplikované.

V první části se autorka věnuje zevrubnému vymezení problémové domény, konceptu komplexního zpracování událostí CEP a jeho využití pro vysokofrekvenční časové řady dat, pojednává o využití vzorů CEP a pravidel pro zpracování v reálném čase. Zdůvodňuje výběr zdroje dat (FOREX největší mimoburzovní trh na světě, na kterém se obchoduje s měnami a kde obchodování probíhá pomocí elektronické sítě a v reálném čase). Autorka dále stanovuje pravidla pro real time zpracování, která pak v aplikační části naplňuje a výsledky hodnotí (využitím platformy ESPER).

V Kapitole 4 vymezuje oblast podpory rozhodování v CEP technologii, vazbu na SOA a systémy pro podporu rozhodování, jejichž modul bude vyvinut v rámci realizace. Kladně hodnotím důraz na podniková pravidla a jejich správnou jednoznačně a interpretovatelnou formalizaci a viditelnou spojitost s událostmi. Výběr konceptu rozhodovací tabulky jako výchozí formalizace pravidel je promyšleným základem vlastního a původního navrženého postupu, tato volba mohla být podrobněji vysvětlena.

V druhé části práce, Kapitole 5 a 6, které považuji za stěžejní a realizační, je postupně budován a prezentován vlastní originální přístup a metoda. Autorka srozumitelně a systémově vysvětluje svoji metodu formalizace pravidel včetně ukázky řešení pro vysokofrekvenční finanční historická data z FOREXu. Kladně hodnotím využití myšlenky oddělení řízení pravidel od spouštění pravidel pomocí tvorby sad pravidel a maticového přístupu. Jde o nový pohled na samotnou formalizaci pravidel, jejíž síla je v jednoduchosti použití a udržovatelnosti sady pravidel.

V této části je rovněž charakterizován nástroj ESPER (open source platforma pro komplexní zpracování událostí CEP) a postupně budován a implementován prototyp nového DSS modulu složeného ze tří logických a podle mého názoru dobře promyšlených částí. Jádrem je formalizace pravidel a použití maticové formální logiky se zahrnutím real time přístupu.

Autorka použila existující platformu CEP a formou nového modulu DSS integrovala svoji metodu formalizace maticovou gramatikou. Reálný výstup disertační práce je zajímavý a původní. Některé formální záležitosti (nevyhraněný přechod od pojmů formalizace resp. model k metodě, která je vlastně samotným cílem práce).

Měření a samotné ověření je popsáno a interpretováno v Kapitole 6. Z mnoha možných parametrů pro vyhodnocování návrhu jsou vybrány rychlost, zpoždění, propustnost a spotřeba paměti při měnících se obchodních strategiích. Nabízí se otázka, proč se neměřily parametry související s obsahem dat, například vhodnost metody pro burzu, rychlost přizpůsobení obchodování změnami pravidel, možnosti optimalizace dat apod. Silnou částí samotného měření pak je srovnání s požadavky pro real time systémy (str. 68-69) vytčenými v první části práce. Zde vidím možnosti využití v jiných oblastech než finančních.

Slabší stránkou práce je podle mého názoru chybějící celkové zhodnocení výsledků v diskuzi. Hodnocení mohlo být zaměřeno například na význam navrženého modelu a jeho vhodnost dalšího využití, na srovnání s jinými přístupy k formalizaci pravidel, možnost aplikace řešení v jiných nástrojích CEP, možnosti rozšíření popsané metody o uživatelskou podporu. V práci postrádám širší diskuzi k řešení a vytvořeným výstupům.

Přesto se domnívám, že výsledek je kvalitní, originální, přináší nový pohled na problematiku a zohledňuje dosažený stupeň poznání v této oblasti. Předložená disertační práce je pečlivá a formálně na dobré úrovni, občasné překlepy nebo výše uvedené edrobné nepřesnosti ji neubírají na kvalitě. Mohu konstatovat, že vytčených cílů bylo dosaženo.

Aktuálnost a původní přínos disertační práce

Podle mého názoru námět a téma disertační práce je z hlediska současného stavu společnosti aktuální. Jde o zajímavé propojení informatiky a obecných principů finanční a statistické oblasti a záměr, přispět k přesnějšímu přístupu k optimalizaci vysokofrekvenčních časových řad, byl naplněn.

Doposud neexistuje, nebo mi není znám, jednoduchý a srozumitelný, ale přitom jasně definovaný, ucelený a univerzálně aplikovatelný rámec pro koncept maticové formální gramatiky zpracované do nástroje pro podporu komplexního zpracování událostí.

Publikační a vědecká činnost autora

Přehled publikační aktivity autora je odpovídající. V seznamu literatury se uvádí celkem 8 publikací (v 7 z nich je hlavní autorkou) v recenzovaných časopisech nebo ve sbornících zahraničních konferencí, kde je jádro práce a jeho význam řádně prezentován.

Doplňující otázky

V rámci rozpravy by se studentka mohla zaměřit na následující témata:

- Z možností aplikačních oblastí, které pokrývá ESPEER jste vybrala oblast burzy. Proč? Jak by podle vás byl váš formalismus vhodný pro řízení podnikových procesů?
- Jak by se úloha a metoda zpracování změnila použitím rozhodovacího stromu (místo matice)?
- Pro jaký typ uživatele je určena vaše metoda? Budou klasičtí burzovní bookmakeři schopni tuto metodu použít? Je možné ji zjednodušit a popsat uživatelsky přátelštěji? Jak by vypadalo uživatelské rozhraní?

Na základě výše uvedených skutečností konstatuji, že disertační práce studentky Ing. Evy Zámečnickové odpovídá požadavkům a po odborné rozpravě doporučuji udělení akademického titulu Ph. D. v oboru Výpočetní technika a informatika.

V Brně dne 14. ledna 2017

doc. Ing. Ivana Rábová Ph.D.

oponentka disertační práce