

Edícia výskumných textov
informatiky a informačných technológií

Štúdie vybraných tém softvérového inžinierstva (2)

**Pokročilé metódy navrhovania programových systémov
Pokročilé metódy získavania, vyhľadávania,
reprezentácie a prezentácie informácie**

Kniha vznikla a bola vydaná s finančnou podporou projektu:

Projekt JPD 3 2004/1-022: Podpora vzdelávania mladých vedeckých pracovníkov
s cieľom vychovať tvorivých expertov – profesionálov informatikov –
pre modernú spoločnosť založenú na vedomostiach



*Európsky sociálny fond pomáha rozvíjať zamestnanosť
podporovaním zamestnateľnosti, obchodného ducha,
rovnakých príležitostí a investovaním
do ľudských zdrojov.*

Mária Bieliková,
Pavol Návrat
a kol.

Štúdie vybraných tém softvérového inžinierstva **2**

Pokročilé metódy navrhovania
programových systémov
Pokročilé metódy získavania, vyhľadávania,
reprezentácie a prezentácie informácie



Slovenská technická univerzita v Bratislave

Fakulta informatiky a informačných technológií

Edícia výskumných textov informatiky a informačných technológií

Štúdie vybraných tém softvérového inžinierstva (2)

Pokročilé metódy navrhovania programových systémov
Pokročilé metódy získavania, vyhľadávania, reprezentácie
a prezentácie informácie

Autorský kolektív:

Mária Bieliková
Pavol Návrat
Anton Andrejko
Anna Bou Ezzeddine
György Frivolt
Jaroslav Jakubík
Martin Kiselkov
Lubomír Majtás
Ján Máté
Ivan Polický
Viliam Solčány
Martin Šechný
Miroslav Vnuk

Fakulta informatiky a informačných technológií
Slovenská technická univerzita v Bratislave
Ilkovičova 3
842 16 Bratislava

© Mária Bieliková, Pavol Návrat a kol., 2006

Text Design & Composition: Mária Bieliková
Copy Editor: Anton Andrejko, Martin Kiselkov
Cover Designer: Peter Kaminský

Vydala Slovenská technická univerzita v Bratislave
vo Vydavateľstve STU, Bratislava, Vazovova 5.

ISBN

PREDHOVOR

Publikácia, ktorá sa vám dostala do rúk, je už druhou v poradí, ktoré sme v relatívne krátkom čase dokázali pripraviť v *Edícii výskumných textov informatiky a informačných technológií* na témy z oblasti programových a informačných systémov. Obe *Štúdie vybraných tém softvérového inžinierstva* sa venujú sa dvom ťažiskovým okruhom. Prvým okruhom sú pokročilé metódy navrhovania programových systémov. Druhým okruhom sú pokročilé metódy získavania, vyhľadávania, reprezentácie a prezentácie informácií. Voľba oboch okruhov tém nebola náhodná. Všetky témy sú aktuálnymi témami súčasného výskumu v oblasti programových a informačných systémov. Ako také sú predmetom záujmu a štúdia výskumných študentov, t.j. najmä študentov doktorandského štúdia. Oni sú nielen prvými čitateľmi *Štúdií*, vybraní doktorandi sú aj autormi jednotlivých častí v oboch publikáciách.

Prvé *Štúdie vybraných tém softvérového inžinierstva* sa v rámci uvedených okruhov zaoberali dvomi ťažiskovými témami. Prvou témou bola analýza návrhových vzorov, ktoré predstavujú jednu z kľúčových oblastí vyvíjajúcej sa disciplíny softvérového inžinierstva. Druhá časť obsahovala päť štúdií z vybraných tém programových a informačných systémov, ktoré diskutujú a analyzujú otvorené vedecké problémy v predmetnej oblasti aj v spojitosti so spracovaním informácií na internete.

Obdobný postup ako pri prvých *Štúdiách* sme zvolili aj pri tomto zväzku. Vznikol na základe seminárov študentov doktorandského štúdia študijného programu programové systémy v odbore softvérové inžinierstvo. Seminára podporil projekt Európskych štrukturálnych fondov, ktorého hlavným cieľom je podpora vzdelávania prostredníctvom motivačných nástrojov pre doktorandov a zvyšovaním kvality vzdelávania v treťom stupni vysokoškolského štúdia v oblasti informatiky a informačných technológií.

Informatika a informačné technológie sú kľúčovým prvkom budovania modernej spoločnosti „založenej na vedomostiach“, ako je dnes módne vraviť. Mladí talentovaní absolventi druhého stupňa vysokoškolského štúdia v oblasti informatiky alebo príbuzných oblastiach majú v súčasnosti veľké možnosti uplatnenia sa v praxi. Informačná spoločnosť však potrebuje aj špecializovaných odborníkov a vedeckých pracovníkov s ukončeným tretím stupňom vysokoškolského štúdia v študijných odboroch skupiny informatických vied, informačných a komunikačných technológií tak, aby bolo možné budovať ekonomiku založenú na najnovších vedeckých poznatkoch. V širšom kontexte ide o rozvoj spoločnosti (ak chcete, založenej na vedomostiach), nielen ekonomiky, schopnej vyrovnávať sa so zložitými výzvami, ktoré pred ňou stoja. S tým súvisí potreba profesionálov v oblasti uchovávanía, spracúvania a prezentácie informácií v bohatej palete reprezentácií ako základného prvku informačnej spoločnosti.

S rozvojom informatiky a informačných technológií a s posunom spoločnosti k informačnej spoločnosti, resp. spoločnosti založenej na vedomostiach, vzniká potreba

vychovávať odborníkov v špecializovaných oblastiach. Seminára, ktoré sa uskutočňujú na Fakulte informatiky a informačných technológií Slovenskej technickej univerzity v Bratislave v rámci doktorandského štúdia a podporené projektom, sa zameriavajú na rôzne oblasti programových a informačných systémov. Zatiaľ čo v prvom zväzku *Štúdií* sme podchytili seminár venovaný návrhovému vzorom, v tomto zväzku sme spracovali témy seminára, venovaného webovej inteligencii. Je to veľmi nová oblasť, možno bližšia viac informačným systémom, avšak v každom prípade potenciálne veľmi užitočná pre každého študenta softvérového inžinierstva..

Našou ambíciou bolo sprístupniť záujemcom o softvérové inžinierstvo vybrané témy a tým zdieľať výsledky seminárov a tvorivého prístupu študentov k jednotlivým témam v rámci diskusií. Výskumné texty v tejto publikácii sú vhodné aj pre študentov ďalších študijných programov v odboroch ako napr. informatika, aplikovaná informatika, informačné systémy, či umelá inteligencia a to v študijných programoch uskutočňovaných na Slovenskej technickej univerzite v Bratislave a aj na iných univerzitách.

Publikácia pozostáva z dvoch dielov. V prvom (Diel 1: Webová inteligencia) sa sústreďujeme na analýzu rôznych aspektov toho, čo sa začalo nazývať webová inteligencia. Druhý (Diel 2: Vybrané témy programových a informačných systémov) obsahuje štyri štúdie, ktoré diskutujú a analyzujú vybrané otvorené vedecké problémy z dynamicky sa rozvíjajúcej oblasti programových systémov so špeciálnym dôrazom na programové informačné systémy aj v spojitosti s internetom.

Diel 1: Webová inteligencia

Čo je to webová inteligencia? Ide o nový pojem. Jeho obsah sa vytvára práve prebiehajúcim výskumom. Výskum sa dotýka širokého spektra otázok, súvisiacich s vývojom ďalšej generácie webu. Je to nový smer vo výskume a vývoji, v ktorom sa skúma, akú rolu alebo praktické dôsledky môže zohrať umelá inteligencia a pokročilé informačné technológie vo vývoji systémov, služieb a prostredí opierajúcich sa o web. Takto to aspoň vidia Ning Zhong, Jiming Liu a Yiyu Yao, ktorí sa podujali zostaviť vôbec prvú (viacautorskú) vedeckú monografiu na túto tému, ktorú vydal v roku 2003 Springer. Knižka sa snaží poskytnúť podrobný úvod a systematický prehľad tejto novej oblasti výskumu. Ponúka štúdie súčasného stavu výskumu jednotlivých problémov. Zaoberá sa tiež niektorými aplikačnými aspektami.

Práve toto boli hlavné dôvody, pre ktoré sme sa rozhodli zamerať doktorandský seminár na jar 2006 na webovú inteligenciu v takom chápaní, v akom ho prezentuje uvedená knižka. Vybrané kapitoly sa stali základom pre referáty, ktoré boli úvodmi pre seminárne diskusie. Seminár v rámci doktorandského štúdia viedol Pavol Návrat. Doktorandi, ktorí referáty predniesli, dopracovali ich textovú podobu potom do výsledného tvaru, ktorý máme možnosť čítať v tomto zväzku.

Každá kapitola je tak výsledkom tvorivej činnosti, ku ktorej prispeli viacerí. Samotný text každej z nich písal ten-ktorý doktorand a jeho autorský prínos treba čo najvýraznejšie zdôrazniť a oceniť. Na seminároch prebiehala diskusia, na ktorej sa zúčastňovala celá skupina doktorandov a ktorá v tom-ktorom prípade ovplyvnila definitívne znenie opisu. Napriek tomu považujeme za korektné, aby sme označili ako jediných autorov jednotlivých opisov doktorandov, ktorí im dali písomnú podobu.

Náš výber tém z webovej inteligencie, ktorý sme zaradili do seminára (a teda aj do tejto knižky), možno rozčleniť do piatich okruhov (kapitol tejto publikácie): webové

agenty (2 témy), dolovanie v dátach na webe (4 témy), vyhľadávanie informácií na webe a manažment znalostí (4 témy), infraštruktúra pre webové inteligentné systémy (2 témy) a inteligencia sociálnych sietí (2 témy).

Autori sa podieľali na jednotlivých kapitolách takto:

- *Webové agenty:*
 - Agentové služby v DAML-S: Martin Kiselkov
 - Návrh scenárov pre sociálne agenty: Ľubomír Majtás
- *Dolovanie v dátach na webe:*
 - Vyhľadávanie informácií porovnávaním webových stránok: Miroslav Vnuk
 - Zisťovanie nepriamych asociácií z dát o používaní webu: Anna Bou Ezzedine
 - Znalostná indukcia obalovačov určených na získavanie informácií z webu: Ľubomír Majtás
 - Dolovanie webových záznamov: Ján Máté
- *Vyhľadávanie informácií na webe a manažment znalostí:*
 - Osobné a zamerané webové pavúky: Martin Kiselkov
 - Reprezentácia, zdieľanie a získavanie znalostí na webe: Miroslav Vnuk
 - Manažment znalostí na webe so sémantikou: Martin Šechný
 - Ontológia – objavovanie taxonomických relácií z webu: Anna Bou Ezzedine
- *Infraštruktúra pre webové inteligentné systémy:*
 - Algoritmické aspekty webových inteligentných systémov: Ivan Polický
 - Predpríprava webových dokumentov na Internete: Ján Máté
- *Inteligencia sociálnych sietí:*
 - Sociálne (spoločenské) siete – od webu k manažmentu znalostí: Martin Šechný
 - Algoritmus usporiadania na určenie miery dobrej povesti alebo relevancie založený na topológii grafu: Ivan Polický

Diel 2: Vybrané témy programových a informačných systémov

Do druhej časti zaraďujeme štyri štúdie, ktoré sa venujú vybraným otvoreným vedeckým problémom, týkajúcim sa programových a informačných systémov. Ide o oblasti, v ktorých prebieha veľmi intenzívny vývoj. Programové systémy sa stávajú systémami, pôsobiacimi v čoraz rôznorodejšom prostredí, vrátane internetu. Stávajú sa súčasťou čoraz komplexnejších systémov – na jednej strane rozsiahlych informačných systémov, na druhej strane systémov, spolu určených technickou platformou, ktorou už dávno nie je len počítač v klasickom slova zmysle, ale aj najrôznejšie vnorené systémy, (tele)komunikačné systémy a pod.

Informačné systémy sa stávajú univerzálnym modelom spôsobov vyhľadávania, získavania, prístupňovania, uchovávanía, odovzdávania, spoločného používania, prezentovania informácií. I keď sa v zásade dá na ne nazerať odhliadnuc od toho, či sú operácie a procesy podporené počítačom alebo nie, čoraz viac sa zväčšuje praktický význam informačných systémov, ktoré sú realizované pomocou programových systémov (a tie samozrejme pomocou počítačových systémov alebo iných technických systémov, zahŕňajúcich počítače). Je to najmä preto, že softvérovo podporené infor-

mačné systémy majú vďaka možnostiam, ktoré poskytuje naprogramovaný počítač, výhody, ktoré sa ručným spracovaním nedajú dosiahnuť. Toto je súčasne aj argumentom pre úzke prepojenie výskumu v oboch oblastiach – ako softvérového inžinierstva, tak aj informačných systémov.

Štúdie sú výsledkom práce doktorandov v rámci ich doktorandského štúdia. Možno nezaškodí pripomenúť, že doktorandské štúdium sa koná pod vedením školiteľa. Na každej štúdií má preto podiel aj príslušný školiteľ. Napriek tomu však považujeme za korektné, aby sme označili ako jediných autorov jednotlivých štúdií doktorandov, ktorí im dali písomnú podobu a ktorí ich predložili a úspešne obhájili ako písomnú časť svojej dizertačnej skúšky.

Autori sa podieľali na jednotlivých kapitolách takto:

- Modelovanie používateľa v aplikáciách opierajúcich sa o web: Anton Andrejko (školiteľ prof. Mária Bieliková)
- Paralelná a distribuovaná simulácia systémov s diskretnými udalosťami: Viliam Solčány (školiteľ: prof. Jiří Šafařík)
- Štruktúry komunit v sieťach: György Frivolt (školiteľ prof. Mária Bieliková)
- Znovupoužitie návrhových vzorov na úrovni zdrojového kódu: Jaroslav Jakubík (školiteľ prof. Pavol Návrat)

Dúfame, že táto knižka poslúži záujemcom o poznanie programových a informačných systémov. Umožňuje spoločne využiť výsledky štúdia v tejto oblasti. Tešíme sa na prípadné odozvy alebo pripomienky.

November 2006,
Bratislava

Mária Bieliková a Pavol Návrat