Kurt Bauknecht Birgit Pröll Hannes Werthner (Eds.)

# E-Commerce and Web Technologies

7th International Conference, EC-Web 2006 Krakow, Poland, September 2006 Proceedings



Kurt Bauknecht Birgit Pröll Hannes Werthner (Eds.)

# E-Commerce and Web Technologies

7th International Conference, EC-Web 2006 Krakow, Poland, September 5-7, 2006 Proceedings







#### Volume Editors

Kurt Bauknecht

University of Zurich, Department of Informatics (IFI)

Winterthurer Strasse 190, 8057 Zurich, Switzerland

E-mail: baukn@ifi.unizh.ch

Birgit Pröll

Johannes Kepler University Linz, Institute for Applied Knowledge Processing (FAW)

Softwarepark Hagenberg, 4232 Hagenberg, Autria

E-mail: bproell@faw.uni-linz.ac.at

Hannes Werthner

Vienna University of Technology

Institute of Software Technology and Interactive Systems, EC-Group

Favoritenstr. 9-11, 1040 Vienna, Austria

E-mail: hannes.werthner@ec.tuwien.ac.at

Library of Congress Control Number: 2006931012

CR Subject Classification (1998): H.4, K.4.4, J.1, K.5, H.3, H.2, K.6.5

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 0302-9743

ISBN-10 3-540-37743-3 Springer Berlin Heidelberg New York ISBN-13 978-3-540-37743-6 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2006 Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India Printed on acid-free paper SPIN: 11823865 06/3142 5 4 3 2 1 0

# Lecture Notes in Computer Science

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Carnegie Mellon University, Pittsburgh, PA, USA

#### **Editorial Board**

David Hutchison

Lancaster University, UK

Takeo Kanade

Josef Kittler University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern
ETH Zurich, Switzerland

John C. Mitchell
Stanford University, CA, USA

Moni Naor
Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz *University of Bern, Switzerland* 

C. Pandu Rangan Indian Institute of Technology, Madras, India

Bernhard Steffen University of Dortmund, Germany

Madhu Sudan

Demetri Terzopoulos University of California, Los Angeles, CA, USA

Massachusetts Institute of Technology, MA, USA

Doug Tygar University of California, Berkeley, CA, USA

Moshe Y. Vardi Rice University, Houston, TX, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

# Lecture Notes in Computer Science

For information about Vols. 1-4035

please contact your bookseller or Springer

- Vol. 4163: H. Bersini, J. Carneiro (Eds.), Artificial Immune Systems. XII, 460 pages. 2006.
- Vol. 4162: R. Královič, P. Urzyczyn (Eds.), Mathematical Foundations of Computer Science 2006. XV, 814 pages. 2006.
- Vol. 4153: N. Zheng, X. Jiang, X. lan (Eds.), Advances in Machine Vision, Image Processing, and Pattern Analysis. XIII, 506 pages. 2006.
- Vol. 4146: J.C. Rajapakse, L. Wong, R. Acharya (Eds.), Pattern Recognition in Bioinformatics. XIV, 186 pages. 2006. (Sublibrary LNBI).
- Vol. 4144: T. Ball, R.B. Jones (Eds.), Computer Aided Verification. XV, 564 pages. 2006.
- Vol. 4139: T. Salakoski, F. Ginter, S. Pyysalo, T. Pahikkala, Advances in Natural Language Processing. XVI, 771 pages. 2006. (Sublibrary LNAI).
- Vol. 4138: X. Cheng, W. Li, T. Znati (Eds.), Wireless Algorithms, Systems, and Applications. XVI, 709 pages. 2006.
- Vol. 4137: C. Baier, H. Hermanns (Eds.), CONCUR 2006 Concurrency Theory. XIII, 525 pages. 2006.
- Vol. 4134: K. Yi (Ed.), Static Analysis. XI, 443 pages. 2006
- Vol. 4133: J. Gratch, M. Young, R. Aylett, D. Ballin, P. Olivier (Eds.), Intelligent Virtual Agents. XIV, 472 pages. 2006. (Sublibrary LNAI).
- Vol. 4130: U. Furbach, N. Shankar (Eds.), Automated Reasoning. XV, 680 pages. 2006. (Sublibrary LNAI).
- Vol. 4129: D. McGookin, S. Brewster (Eds.), Haptic and Audio Interaction Design. XII, 167 pages. 2006.
- Vol. 4128: W.E. Nagel, W.V. Walter, W. Lehner (Eds.), Euro-Par 2006 Parallel Processing. XXXIII, 1221 pages. 2006.
- Vol. 4127: E. Damiani, P. Liu (Eds.), Data and Applications Security XX. X, 319 pages. 2006.
- Vol. 4124: H. de Meer, J.P. G. Sterbenz (Eds.), Self-Organising Systems. XIV, 261 pages. 2006.
- Vol. 4121: A. Biere, C.P. Gomes (Eds.), Theory and Applications of Satisfiability Testing SAT 2006. XII, 438 pages. 2006.
- Vol. 4119: C. Dony, J.L. Knudsen, A. Romanovsky, A. Tripathi (Eds.), Advanced Topics in Exception Handling Components. X, 302 pages. 2006.
- Vol. 4117: C. Dwork (Ed.), Advances in Cryptology Crypto 2006. XIII, 621 pages. 2006.
- Vol. 4115: D.-S. Huang, K. Li, G.W. Irwin (Eds.), Computational Intelligence and Bioinformatics, Part III. XXI, 803 pages. 2006. (Sublibrary LNBI).

- Vol. 4114: D.-S. Huang, K. Li, G.W. Irwin (Eds.), Computational Intelligence, Part II. XXVII, 1337 pages. 2006. (Sublibrary LNAI).
- Vol. 4113: D.-S. Huang, K. Li, G.W. Irwin (Eds.), Intelligent Computing, Part I. XXVII, 1331 pages. 2006.
- Vol. 4112: D.Z. Chen, D. T. Lee (Eds.), Computing and Combinatorics. XIV, 528 pages. 2006.
- Vol. 4111: F.S. de Boer, M.M. Bonsangue, S. Graf, W.-P. de Roever (Eds.), Formal Methods for Components and Objects. VIII, 447 pages. 2006.
- Vol. 4109: D.-Y. Yeung, J.T. Kwok, A. Fred, F. Roli, D. de Ridder (Eds.), Structural, Syntactic, and Statistical Pattern Recognition. XXI, 939 pages. 2006.
- Vol. 4108: J.M. Borwein, W.M. Farmer (Eds.), Mathematical Knowledge Management. VIII, 295 pages. 2006. (Sublibrary LNAI).
- Vol. 4106: T.R. Roth-Berghofer, M.H. Göker, H. A. Güvenir (Eds.), Advances in Case-Based Reasoning. XIV, 566 pages. 2006. (Sublibrary LNAI).
- Vol. 4104: T. Kunz, S.S. Ravi (Eds.), Ad-Hoc, Mobile, and Wireless Networks. XII, 474 pages. 2006.
- Vol. 4099: Q. Yang, G. Webb (Eds.), PRICAI 2006: Trends in Artificial Intelligence. XXVIII, 1263 pages. 2006. (Sublibrary LNAI).
- Vol. 4098: F. Pfenning (Ed.), Term Rewriting and Applications. XIII, 415 pages. 2006.
- Vol. 4097: X. Zhou, O. Sokolsky, L. Yan, E.-S. Jung, Z. Shao, Y. Mu, D.C. Lee, D. Kim, Y.-S. Jeong, C.-Z. Xu (Eds.), Emerging Directions in Embedded and Ubiquitous Computing. XXVII, 1034 pages. 2006.
- Vol. 4096: E. Sha, S.-K. Han, C.-Z. Xu, M.H. Kim, L.T. Yang, B. Xiao (Eds.), Embedded and Ubiquitous Computing. XXIV, 1170 pages. 2006.
- Vol. 4094: O. H. Ibarra, H.-C. Yen (Eds.), Implementation and Application of Automata. XIII, 291 pages. 2006.
- Vol. 4093: X. Li, O.R. Zaïane, Z. Li (Eds.), Advanced Data Mining and Applications. XXI, 1110 pages. 2006. (Sublibrary LNAI).
- Vol. 4092: J. Lang, F. Lin, J. Wang (Eds.), Knowledge Science, Engineering and Management. XV, 664 pages. 2006. (Sublibrary LNAI).
- Vol. 4091: G.-Z. Yang, T. Jiang, D. Shen, L. Gu, J. Yang (Eds.), Medical Imaging and Augmented Reality. XIII, 399 pages. 2006.
- Vol. 4090: S. Spaccapietra, K. Aberer, P. Cudré-Mauroux (Eds.), Journal on Data Semantics VI. XI, 211 pages. 2006.
- Vol. 4089: W. Löwe, M. Südholt (Eds.), Software Composition. X, 339 pages. 2006.

Vol. 4088: Z.-Z. Shi, R. Sadananda (Eds.), Agent Computing and Multi-Agent Systems. XVII, 827 pages. 2006. (Sublibrary LNAI).

Vol. 4085: J. Misra, T. Nipkow, E. Sekerinski (Eds.), FM 2006: Formal Methods. XV, 620 pages. 2006.

Vol. 4082: K. Bauknecht, B. Pröll, H. Werthner (Eds.), E-Commerce and Web Technologies. XIII, 243 pages. 2006.

Vol. 4079: S. Etalle, M. Truszczyński (Eds.), Logic Programming. XIV, 474 pages. 2006.

Vol. 4077: M.-S. Kim, K. Shimada (Eds.), Geometric Modeling and Processing - GMP 2006. XVI, 696 pages. 2006.

Vol. 4076: F. Hess, S. Pauli, M. Pohst (Eds.), Algorithmic Number Theory. X, 599 pages. 2006.

Vol. 4075: U. Leser, F. Naumann, B. Eckman (Eds.), Data Integration in the Life Sciences. XI, 298 pages. 2006. (Sublibrary LNBI).

Vol. 4074: M. Burmester, A. Yasinsac (Eds.), Secure Mobile Ad-hoc Networks and Sensors. X, 193 pages. 2006.

Vol. 4073: A. Butz, B. Fisher, A. Krüger, P. Olivier (Eds.), Smart Graphics. XI, 263 pages. 2006.

Vol. 4072: M. Harders, G. Székely (Eds.), Biomedical Simulation. XI, 216 pages. 2006.

Vol. 4071: H. Sundaram, M. Naphade, J.R. Smith, Y. Rui (Eds.), Image and Video Retrieval. XII, 547 pages. 2006.

Vol. 4070: C. Priami, X. Hu, Y. Pan, T.Y. Lin (Eds.), Transactions on Computational Systems Biology V. IX, 129 pages. 2006. (Sublibrary LNBI).

Vol. 4069: F.J. Perales, R.B. Fisher (Eds.), Articulated Motion and Deformable Objects. XV, 526 pages. 2006.

Vol. 4068: H. Schärfe, P. Hitzler, P. Øhrstrøm (Eds.), Conceptual Structures: Inspiration and Application. XI, 455 pages. 2006. (Sublibrary LNAI).

Vol. 4067: D. Thomas (Ed.), ECOOP 2006 – Object-Oriented Programming. XIV, 527 pages. 2006.

Vol. 4066: A. Rensink, J. Warmer (Eds.), Model Driven Architecture – Foundations and Applications. XII, 392 pages, 2006.

Vol. 4065: P. Perner (Ed.), Advances in Data Mining. XI, 592 pages. 2006. (Sublibrary LNAI).

Vol. 4064: R. Büschkes, P. Laskov (Eds.), Detection of Intrusions and Malware & Vulnerability Assessment. X, 195 pages. 2006.

Vol. 4063: I. Gorton, G.T. Heineman, I. Crnkovic, H.W. Schmidt, J.A. Stafford, C.A. Szyperski, K. Wallnau (Eds.), Component-Based Software Engineering. XI, 394 pages. 2006.

Vol. 4062: G. Wang, J.F. Peters, A. Skowron, Y. Yao (Eds.), Rough Sets and Knowledge Technology. XX, 810 pages. 2006. (Sublibrary LNAI).

Vol. 4061: K. Miesenberger, J. Klaus, W. Zagler, A. Karshmer (Eds.), Computers Helping People with Special Needs. XXIX, 1356 pages. 2006.

Vol. 4060: K. Futatsugi, J.-P. Jouannaud, J. Meseguer (Eds.), Algebra, Meaning, and Computation. XXXVIII, 643 pages. 2006.

Vol. 4059: L. Arge, R. Freivalds (Eds.), Algorithm Theory – SWAT 2006. XII, 436 pages. 2006.

Vol. 4058: L.M. Batten, R. Safavi-Naini (Eds.), Information Security and Privacy. XII, 446 pages. 2006.

Vol. 4057: J.P.W. Pluim, B. Likar, F.A. Gerritsen (Eds.), Biomedical Image Registration. XII, 324 pages. 2006.

Vol. 4056: P. Flocchini, L. Gasieniec (Eds.), Structural Information and Communication Complexity. X, 357 pages. 2006.

Vol. 4055: J. Lee, J. Shim, S.-g. Lee, C. Bussler, S. Shim (Eds.), Data Engineering Issues in E-Commerce and Services. IX, 290 pages. 2006.

Vol. 4054: A. Horváth, M. Telek (Eds.), Formal Methods and Stochastic Models for Performance Evaluation. VIII, 239 pages. 2006.

Vol. 4053: M. Ikeda, K.D. Ashley, T.-W. Chan (Eds.), Intelligent Tutoring Systems. XXVI, 821 pages. 2006.

Vol. 4052: M. Bugliesi, B. Preneel, V. Sassone, I. Wegener (Eds.), Automata, Languages and Programming, Part II. XXIV, 603 pages. 2006.

Vol. 4051: M. Bugliesi, B. Preneel, V. Sassone, I. Wegener (Eds.), Automata, Languages and Programming, Part I. XXIII, 729 pages. 2006.

Vol. 4049: S. Parsons, N. Maudet, P. Moraitis, I. Rahwan (Eds.), Argumentation in Multi-Agent Systems. XIV, 313 pages. 2006. (Sublibrary LNAI).

Vol. 4048: L. Goble, J.-J.C.. Meyer (Eds.), Deontic Logic and Artificial Normative Systems. X, 273 pages. 2006. (Sublibrary LNAI).

Vol. 4047: M. Robshaw (Ed.), Fast Software Encryption. XI, 434 pages. 2006.

Vol. 4046: S.M. Astley, M. Brady, C. Rose, R. Zwiggelaar (Eds.), Digital Mammography. XVI, 654 pages. 2006.

Vol. 4045: D. Barker-Plummer, R. Cox, N. Swoboda (Eds.), Diagrammatic Representation and Inference. XII, 301 pages. 2006. (Sublibrary LNAI).

Vol. 4044: P. Abrahamsson, M. Marchesi, G. Succi (Eds.), Extreme Programming and Agile Processes in Software Engineering. XII, 230 pages. 2006.

Vol. 4043: A.S. Atzeni, A. Lioy (Eds.), Public Key Infrastructure. XI, 261 pages. 2006.

Vol. 4042: D. Bell, J. Hong (Eds.), Flexible and Efficient Information Handling. XVI, 296 pages. 2006.

Vol. 4041: S.-W. Cheng, C.K. Poon (Eds.), Algorithmic Aspects in Information and Management. XI, 395 pages. 2006.

Vol. 4040: R. Reulke, U. Eckardt, B. Flach, U. Knauer, K. Polthier (Eds.), Combinatorial Image Analysis. XII, 482 pages. 2006.

Vol. 4039: M. Morisio (Ed.), Reuse of Off-the-Shelf Components. XIII, 444 pages. 2006.

Vol. 4038: P. Ciancarini, H. Wiklicky (Eds.), Coordination Models and Languages. VIII, 299 pages. 2006.

Vol. 4037: R. Gorrieri, H. Wehrheim (Eds.), Formal Methods for Open Object-Based Distributed Systems. XVII, 474 pages. 2006.

Vol. 4036: O. 怕. Ibarra, Z. Dang (Eds.), Developments in Language Theory. XII, 456 pages. 2006.

辛385.02元

#### **Preface**

We welcome you to the 7th International Conference on E-commerce and Web Technologies (EC-Web 2006) held in Krakow, Poland, in conjunction with DEXA 2006. This conference was organized for the first time in Greenwich, UK, in 2000, and it has been able to attract an increasing number of participants and interest, reflecting the progress made in the field. As in the previous years, EC-Web 2006 served as a forum bringing together researchers from academia and practitioners from industry to discuss the current state of the art in e-commerce and Web technologies. Inspiration and new ideas emerged from intensive discussions that took place during the keynote address, the formal sessions and the social events.

The conference attracted 88 paper submissions and each paper was reviewed by three Program Committee members. The Program Committee selected 24 papers for presentation and publication (an acceptance and publication rate of 27%). We have to confess that this task was not that easy due to the high quality of the submitted papers.

We would like to express our thanks to our colleagues who helped put together the technical program: the Program Committee members and external reviewers for their timely and rigorous reviews of the papers, and the Organizing Committee for their help in the administrative work and support. We owe special thanks to Gabriela Wagner for her helping hand concerning the administrative and organizational tasks of this conference.

Finally, we would like to thank all the authors who have submitted papers, authors who presented papers, and the participants who made this conference an intellectually stimulating event.

We hope that all attendees enjoyed the hospitality of Krakow and the conference.

August 2006

Birgit Pröll (Johannes Kepler University of Linz, Austria) Hannes Werthner (Vienna University of Technology, Austria) Program Chairs EC-Web 2006

#### **Organization**

#### **Program Committee**

Marco Aiello, University of Trento, Italy Sourav S. Bhowmick, Nanyang Technological University, Singapore Martin Bichler, Technical University Munich, Germany Susanne Boll, University of Oldenburg, Germany Stephane Bressan, National University of Singapore, Singapore Tanya Castleman, Deakin University, Australia Wojciech Cellary, The Poznan University of Economics, Poland Jen-Yao Chung, IBM T.J. Watson Research Center, USA Alfredo Cuzzocrea, University of Calabria, Italy Eduardo Fernandez, Florida Atlantic University, USA Elena Ferrari, University of Insubria at Como, Italy Farshad Fotouhi, Wayne State University, USA Karl A. Fröschl, Electronic Commerce Competence Center, Austria Yongjian Fu, Cleveland State University, USA Stephane Gagnon, New Jersey Institute of Technology, USA Fausto Giunchiglia, University of Trento, Italy Chanan Glezer, Ben Gurion University, Israel Thomas Hess, LMU Munich, Germany Yigal Hoffner, Switzerland Christian Huemer, University of Vienna, Austria Gregory E. Kersten, Concordia University Montreal, Canada Hiroyuki Kitagawa, University of Tsukuba, Japan Gabriele Kotsis, Johannes Kepler University Linz, Austria Alberto Laender, Federal University of Minas Gerais, Brazil Juhnyoung Lee, IBM T. J. Watson Research Center, USA Leszek Lilien, Western Michigan University, USA Ee-Peng Lim, Nanyang Technological University, Singapore Huan Liu, Arizona State University, USA Heiko Ludwig, IBM T. J. Watson Research Center, USA Sanjay Kumar Madria, University of Missouri-Rolla, USA Bamshad Mobasher, DePaul University, USA Natwar Modani, IBM India Research Lab, India Mukesh Mohania, IBM India Research Lab, India Guenter Mueller, University of Freiburg, Germany Dirk Neumann, University of Karlsruhe, Germany Gustaf Neumann, Vienna University of Economics and BA, Austria Wee Keong Ng, Nanyang Technological University, Singapore Rolf Oppliger, eSECURITY Technologies, Switzerland Oscar Pastor, Valencia University of Technology, Spain Guenther Pernul, University of Regensburg, Germany

Evangelia Pitoura, University of Ioannina, Greece Ivana Podnar, EPFL, Switzerland Giuseppe Psaila, University of Bergamo, Italy Gerald Quirchmayr, University of Vienna, Austria Indrakshi Ray, Colorado State University, USA Werner Retschitzegger, Johannes Kepler University Linz, Austria Tomas Sabol, Technical University of Kosice, Slovakia Nandlal L. Sarda, Indian Institute of Technology Bombay, India Steffen Staab, University of Koblenz, Germany Michael Stroebel, BMW Group, Germany Roger M. Tagg, University of South Australia, Australia Kian-Lee Tan, National University of Singapore, Singapore Stephanie Teufel, University of Fribourg, Switzerland Bruce H. Thomas, University of South Australia, Australia A Min Tioa, Technical University of Vienna, Austria Aphrodite Tsalgatidou, University of Athens, Greece Krishnamurthy Vidyasankar, Memorial University of Newfoundland, Canada Hans Weigand, Tilburg University, The Netherlands Christof Weinhardt, University of Karlsruhe, Germany Janusz Wielki, Technical University of Opole, Poland

#### **External Reviewers**

Nitin Agarwal, Arizona State University, USA George Athanasopoulos, University of Athens, Greece Michael Borovicka, University of Innsbruck, Austria Peter Butka, Technical University of Kosice, Slovakia Karol Furdik, Intersoft, a.s., Slovakia Woiciech Galuba, EPFL, Switzerland Jörg Gilberg, University of Regensburg, Germany Christoph Grün, Vienna University of Technology, Austria Fabius Klemm, EPFL, Switzerland Jan Kolter, University of Regensburg, Germany Marian Mach, Technical University of Kosice, Slovakia Patrick Sinclair Merten, University of Fribourg, Switzerland Sai Moturu, Arizona State University, USA Björn Muschall, University of Regensburg, Germany Michael Pantazoglou, University of Athens, Greece Marek Paralic, Technical University of Kosice, Slovakia Lance Parsons, Arizona State University, USA Vicente Pelechano, Valencia University of Technology, Spain Gonzalo Rojas, Valencia University of Technology, Spain Marta Ruiz, Valencia University of Technology, Spain Jarogniew Rykowski, Poznan University of Economics, Poland Ali Salehi, EPFL, Switzerland Martin Steinert, University of Fribourg, Switzerland

Sergiusz Strykowski, Poznan University of Economics, Poland Lei Tang, Arizona State University, USA Victoria Torres, Valencia University of Technology, Spain Pedro Valderas, Valencia University of Technology, Spain Le-Hung Vu, EPFL, Switzerland Daniela Wanner, University of Fribourg, Switzerland Marco Zapletal, Vienna University of Technology, Austria Zheng Zhao, Arizona State University, USA

# **Table of Contents**

Recommender System 1

Map-Based Recommendation of Hyperlinked Document Collections Mieczysław A. Kłopotek, Sławomir T. Wierzchoń, Krzysztof Ciesielski, Michał Dramiński, Dariusz Czerski	1
Web User Segmentation Based on a Mixture of Factor Analyzers Yanzan Kevin Zhou, Bamshad Mobasher	11
A Hybrid Similarity Concept for Browsing Semi-structured Product Items	21
Recommender Systems 2	
A Preference-Based Recommender System	31
Enhanced Prediction Algorithm for Item-Based Collaborative Filtering Recommendation	41
Persuasive Online-Selling in Quality and Taste Domains	51
Business Process / Design Aspects	
Proviado – Personalized and Configurable Visualizations of Business Processes	61
Service-Oriented Data and Process Models for Personalization and Collaboration in e-Business	72
A Framework for Raising Collaboration Levels on the Internet	82

Designing Volatile Functionality in E-Commerce Web Applications Gustavo Rossi, Andres Nieto, Luciano Mengoni, Liliana Nuño Silva	92
Mobile Commerce	
Design of Ubiquitous Referral Marketing: A Business Model and Method	102
Pre-service and Post-transcoding Schema for an Adaptive PC to Mobile Web Contents Transcoding System	112
Context-Aware Recommendation Service Using Multi-leveled Information in Mobile Commerce	122
Security and E-Payment	
Attribute-Based Authentication and Authorisation Infrastructures for E-Commerce Providers	132
Seamlessness and Privacy Enhanced Ubiquitous Payment	142
CerTicket Solution: Safe Home-Ticketing Through Internet Francisco Rico, Jordi Forga, Emilio Sanvicente, Luis de la Cruz	152
Web Services Computing / Semantic Web	
Efficient Invocation of Web Services Using Intensional Results	162
On Distributed Service Selection for QoS Driven Service Composition	173
RLinda: A Petri Net Based Implementation of the Linda Coordination Paradigm for Web Services Interactions	183

Table of Contents

XIII

# Map-Based Recommendation of Hyperlinked Document Collections

Mieczysław A. Kłopotek, Sławomir T. Wierzchoń, Krzysztof Ciesielski, Michał Dramiński, and Dariusz Czerski

Institute of Computer Science, Polish Academy of Sciences, ul. Ordona 21, 01-237 Warszawa, Poland {kciesiel, klopotek, mdramins, stw, dcz}@ipipan.waw.pl

**Abstract.** The increasing number of documents returned by search engines for typical requests makes it necessary to look for new methods of representation of the search results.

In this paper, we discuss the possibility to exploit incremental, navigational maps based both on page content, hyperlinks connecting similar pages and ranking algorithms (such as HITS, SALSA, PHITS and PageRank) in order to build visual recommender system. Such system would have an immediate impact on business information management (e.g. CRM and marketing, consulting, education and training) and is a major step on the way to information personalization.

#### 1 Introduction

Recommender systems became indispensable part of modern e-business, especially using the electronic medium of the Internet. It is claimed that even 20% of clients may be encouraged to purchase a good by the most successful recommender systems. Recommenders are applied in advertisement of books (Amazon), CD's (MediaUnbound, MoodLogic, CDNow, SongExplorer), films (Reel.com Movie Matcher, MovieFinder.com Mach Maker), cosmetics (Drugstore.com) and other. But also recommender systems are applied when advice is sought by firms when selecting training courses for employees, job search for unemployed etc.

Recommender systems are not only applied to increase sales, but also to expand cross-selling, attraction of new clients, extending of trust of existent clients, to overcome the barriers of mass advertisement through high personalization of offers. The intense research did not, however, overcome many important limitations, like sparseness of data [21], scalability and real time action, detail level of available data, feedback to the sales firm [22], protection against preference manipulation, visualization of recommendation acceptability reasons [15], client modeling, system evaluation, creation of distributed systems [3], etc.

In our on-going research effort, we target at recommender systems capable of overcoming the limitations of present-day systems with respect to problems of rare data in recommendation, scalability and visualization of recommendation

K. Bauknecht et al. (Eds.): EC-Web 2006, LNCS 4082, pp. 1-10, 2006.

for the purpose of proper explanation and justification of recommendation. Results of this research will surely lead to practical guidelines for construction of commercial recommender systems, where above mentioned problems are crucial.

We have created a full-fledged search engine BEATCA for small collections of documents (up to several millions) capable of representing on-line replies to queries in graphical form on a document map. We extended WebSOM's goals by a multilingual approach, new forms of geometrical representation and we experimented also with various modifications to the clustering process itself [17,18] . The crucial issue for understanding the 2D map by the user is the clustering of its contents and appropriate labeling of the clustered map areas.

Several important issues need to be resolved in order to present the user with an understandable map of documents. The first issue is the way of document clustering. In the domains, like e.g. legal documents, where the concepts are not sharply separated, a fuzzy-set theoretic approach to clustering appears to be a promising one. The other one is the issue of initialization of topical maps. Our experiments showed that the random initialization performed in the original WebSOM may not lead to appearance of meaningful structure of the map. Therefore, we proposed several methods for topical map initialization, based on SVD, PHITS, PLSA and Bayesian network techniques.

In this paper, before we report on our current state of research effort starting with Sect. 4, we briefly present an overview of recommender system concepts (Sect. 2) and our concept of an integrated recommender system (Sect. 3).

# 2 Recommender Systems Overview

Intelligent agent (IA) is a user's assistant or recommender system based on machine learning and data mining techniques. Construction of IA uses the following paradigm taken from the ordinary life: "people uses helpful information without any fixed plans". People do not need and cannot describe their own work in terms of coefficients and classification. People just operate and know what they are interesting in, or what they want when they see it.

Recommender system simulates some social behavior. No one has unlimited knowledge and such knowledge is not necessary on daily basis. However, in some decision problems we have to go into details of specific, narrow knowledge. Sometimes there is a possibility to use advice from experienced person (expert) in a given area. Recommender systems try to help user in such situation by using knowledge collected in specified discipline and watching decisions made by other users decisions in the similar case. These systems have been built as a help in decision process for people, but also for multi-agent systems and generally speaking for systems consisting of objects that have limited knowledge about environment. Recommender system uses knowledge about passive objects to recommend next (somehow similar) item to active objects. For example, recommender system can recommend next web page or article in Internet shop (passive object) that user (active object) is probably looking for.

Recommender systems may be classified along the following criteria: amount and type of data that come from active object, amount and type of required data

about community of active objects, method of recommendation, result type of recommendation, way of delivering recommendation to the active object and the degree of personalization (adaptivity to active object characteristic). More detailed classification and examples of commercial recommender systems can be found in [22].

Methods of recommendation in early systems were based mostly on the following approaches: recommendations based on searching, categories, clustering, association rules or classifiers. Finally, evolution of recommender systems has led to two major approaches in construction of IA:

- 1. Content-based approach. System creates users profiles by analyzing their operations and recommends documents that are compatible with these profiles.
- 2. Collaborative approach collaborative or social filtering [12]. System focuses on a group of users.

The first approach, rooted in the tradition of information processing, is applicable if the system deals with text only. The system seeks information similar to that preferred by the user. If a user is interested in some knowledge areas (represented by documents described by some keywords or phrases) then the recommender looks for documents with similar content to already articulated. The basic problem here is to capture all specific aspects of a document content (e.g. in disciplines such as music, film, computer-related issues etc.). Even restricting recommendations to text documents only, most representations are able to cover only some aspects of document content, which results in weak quality of presented recommendations.

The second approach, called also social learning, relies on exploiting reactions of other users to the same object (e.g. a course, educational path, a film, etc.). The system looks for users with similar interests, capabilities etc. and recommends them information or items they are searching for.

This approach allows for posing questions like "show me information I have never seen but it turned interesting to people like me". Personalized information is provided in an iterative process where information is presented and user is asked to rank it, what allows to determine his/her profile. This profile is next used to locate other users with similar interests, in order to identify groups with similar interests.

Instead of calculating similarity between documents, this method determines degree of membership to a group (for example based on surveys). In contrast to first approach, it does not require analysis of document content, what means that document with arbitrary content could be presented to the user, with the same probability. Each document is assigned with identifier and a degree of membership to a group.

This approach is characterized by two features: first of all the document relevance is determined in the context of the group and not of a single user. Second, evaluation of the document is subjective. Hence one can handle complex and heterogeneous evaluation schemas.

#### 3 Integrated Recommendations

Existing recommender systems, based on a paradigm of content-based filtering as well as those based on collective filtering principle, do not take into consideration possible synergic effects. Such effects emerge when:

- both methodologies are merged,
- system is able to model joint, integrated recommendation of passive and active objects (i.e. clients and products), and not only passive objects pointed by active ones,
- recommendations are based on visual system, which helps to explain and justify a recommendation.

Application of joint methodology is possible if available data contain information on recommended objects as well as relations between recommended and recommending objects. Such information is present, e.g. in WWW documents, where individual html pages have not only textual context, but also hyperlinks between them. From logs saved on a particular host one can obtain so-called click-stream of users surfing from one page to another, and some additional data such as voluntarily filled-in questionnaires. Among other examples are libraries, book stores, or any shop (including e-shops), where products can be described by a set of attributes (e.g. advertisement leaflet) and users can be identified by some ID cards (e.g. loyalty program participation cards). Similarly, for some services (e.g. concerning education or health), both pointed(passive) and pointing(active) objects are described by attributes.

By an *integrated recommendation* we mean recommendation such as "People interested in <characteristics of people> are buying also book <title>" (instead of typical recommendation in form: "People interested in <title> are buying also book <title>"). Thus, integrated recommendation requires that system has an ability to generalize features describing characteristics of active objects (i.e. users or clients).

Recommendation with a visual explanation and justification is a completely new approach, based on creation of two-dimensional, navigational map of objects. Such a map yields a possibility to present an identified area of user's interests together with surrounding context, i.e. main directions of his/her future activities.

#### 4 BEATCA Search Engine

Our first step towards a new model of recommendation system was to create a new-type search engine, based on a document map interface. Our map-based approach to search engine interfacing comprises two important features from the point of view of the target recommendation system: providing an overview over the whole collection of objects, and a very detailed clustering into groups of objects and their immediate (local) contexts.

With a strongly parameterized map creation process, the user of BEATCA can accommodate map generation to his particular needs, or even generate multiple