

WebKDD 2008: 10 years of Knowledge Discovery on the Web Post-Workshop report

Olfa Nasraoui
Computer Engineering and Computer
Science department
Speed School of Engineering
University of Louisville
Louisville KY 40292
olfa.nasraoui@louisville.edu

Myra Spiliopoulou
Otto-von-Guericke-University Magdeburg,
Faculty of Computer Science
D-39016 Magdeburg, Germany
myra@iti.cs.uni-magdeburg.de

Osmar R. Zaiane
Dept. of Computer Science
University of Alberta
Edmonton, Canada
zaiane@cs.ualberta.ca

Jaideep Srivastava
Dept. of Computer Science and
Engineering
University of Minnesota
Minneapolis, MN 55455
srivasta@cs.umn.edu

Bamshad Mobasher
School of Computing
DePaul University
Chicago, IL 60604
mobasher@cs.depaul.edu

ABSTRACT

WebKDD was held in conjunction with the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2008), August 24, 2008, in Henderson/Las Vegas, Nevada. In 2008, WebKDD was held for the tenth time. We report on the contents and outcomes of this anniversary workshop and include a small retrospection on 10 years of knowledge discovery on the Web.

Keywords

Web mining, Web analytics, personalization, recommender systems, query-log mining, communities, attack detection, spam detection

1. INTRODUCTION

WebKDD 2008 was the tenth of a successful series of workshops. Launched in 1999 by Brij Masand and Myra Spiliopoulou, the first workshop of the series WebKDD 1999 invited contributions on “Web usage mining”. In the 10 years that followed, the scope of WebKDD was broadened to cover the emerging KDD topics on the Web: Mining Web usage and content, extracting semantics, analyzing Web query logs and supporting Web search, addressing the challenges of personalization and recommender systems, protecting user privacy and fending off attacks and spam.

For the past nine years, WebKDD workshops have fostered new ideas and research contributions on the hot topics of Web mining. Many workshops have concentrated on one subject area: For example, WebKDD 2000 and WebKDD 2001 focused on e-commerce, WebKDD 2005 stressed the aspect of evolution in Web clickstreams, while WebKDD 2007 was on mining the Social Web and was organized

jointly with the SNAKDD 2007 on social network analysis. The 10th WebKDD workshop invited contributions on all areas of Web mining and further promoted discussions on “lessons learned” and open issues for the years to follow. WebKDD 2008 also featured a best-paper-award, sponsored by Yahoo! Research.

2. TEN YEARS OF WEBKDD

WebKDD 2008 called for retrospection and reflection upon the achievements and the open issues of Web Mining. Some of these achievements are mirrored in the impact of the workshop series itself: In February 2008, when the proposal for the 10th WebKDD workshop was submitted for evaluation, the papers at the past WebKDD workshops had accumulated about 940 citations, including 7 patent citations that cite the workshop as the premier reference on the topic of Web usage analysis.

One day before the workshop opening, on August 22, 2008, the WebKDD 2000 paper “Application of Dimensionality Reduction in Recommender System - A Case Study” [10] counted 299 citations, while an earlier WebKDD 1999 publication, “Data Mining of User Navigation Patterns” [11], was cited 228 times. Although earlier papers have an advantage over more recent ones, and generic contributions tend to be more broadly cited than specialized ones, contributions of more recent years have also seen their share: Among these, the rather specialized work on “Using Rank Propagation and Probabilistic Counting for Link-Based Spam Detection” [12] from WebKDD 2006 was cited 26 times and was one of 25 papers with more than 20 citations.

To bring the recent advances in all those cited areas together, WebKDD 2008 invited contributions on all aspects of user activities in the Web: (1) Mining server log (clickstream) data, (2) mining query log data, (3) mining user activities in social venues and (4) mining user activities in venues that foster the capturing of data semantics, e.g. for building the Semantic Web or for building folksonomies.

To promote reflection on lessons learned and on open issues, WebKDD 2008 included time for discussions at the end of each session, as well as a closing discussion after all papers have been presented.

3. SUBMISSIONS

WebKDD 2008 received 24 submissions. Each paper was peer-reviewed by three members of the Program Committee and 9 papers were accepted for presentation. Among them, one paper was selected for the best paper award, as described below. An extended version of the accepted papers will appear in an LNCS volume in the series “Advances in Web Mining and Web Usage Analysis” that has accommodated the extended WebKDD papers since 2005.

The materials of the workshop, including the program and the slides of the talks can be found on <http://webmining.spd.louisville.edu/webkdd08/program.html>.

4. WORKSHOP

The accepted papers were organized into four sessions. To allow for a deeper content discussion, the sessions were further grouped across the two themes, “Social and Web Graph Mining” and “Web Personalization and User Profiling”. The intensity of the discussions was remarkable: During and after the talks there were well-informed questions and comments from the approximately thirty participants.

4.1 Social and Web Graph Mining

The first paper, “*Detecting Communities via Simultaneous Clustering of Graphs and Folksonomies*” by Akshay Java, Anupam Joshi, and Tim Finin [1], present a simple technique for detecting communities by utilizing both the link structure and folksonomy (or tag) information that is readily available in most social media systems. A simple way to describe their approach is by defining a community as a set of nodes in a graph that link more frequently to within this set than outside it and they share similar tags. The proposed technique is based on the Normalized Cut (NCut) algorithm and can be easily and efficiently implemented. Their method, Simultaneous Cut (SimCut), has the advantage that it can group related tags and cluster the nodes simultaneously.

The second paper, “*I/O efficient computation of First Order Markov Measures for Large and Evolving Graphs*” by Prasanna K Desikan and Jaideep Srivastava [2], addresses the challenges of computing First-order Markov measures, considering PageRank as the example of such a measure. They address two challenging computational scenarios for PageRank: (a) computation for a large single graph at a given time instance and (b) incremental computation for large evolving graphs. They achieve efficiency by reducing the problem size and reducing the number of iterations to compute. For (a), they bin the nodes in different partitions and for the subgraph formed by each of these partitions they use the nature of the first-order Markov model to break down the problem of computation. For (b), they propose a method to accommodate the changed edges and nodes into new partitions and existing partitions and identify the subset of partitions for which re-computation is necessary. For each

identified partition, they use an incremental approach to compute the measure in a scalable manner.

The third paper, “*An Empirical Study of Dynamic Pari-mutuel Markets: Evidence from the Tech Buzz Game*” by Yiling Chen, David Pennock, and Tejaswi Kasturi [3], examines empirical questions related to DPMs. A dynamic pari-mutuel market (DPM) is a hybrid between a continuous double auction (CDA) and a pari-mutuel market. Like a CDA, a DPM incentivizes traders to reveal their information early. Like a pari-mutuel market, a DPM has infinite liquidity, allowing trades at any time. The paper examines the following two questions: Do prices in DPMs predict events of interests? How do traders behave in DPMs? Leveraging a data set from the Yahoo!/O’Reilly Tech Buzz Game, a live system using the DPM, the authors show that prices offer informative forecasts of future event trends. At the agent level, they find that on average human traders outperform robot traders who randomly place trades. Those human traders who are seemingly more rational, buying when the implied market probability is low and selling when it is high, obtain higher profit on average than those who appear less rational. They also examine other aspects of the game, including incentives to manipulate the market.

The fourth paper, “*Query-log mining for detecting polysemy and spam*” by Carlos Castillo, Claudio Corsi, Debora Donato, Paolo Ferragina, and Aristides Gionis [4], exploits user feedback and query-log mining to combat spam and identify query polysemy. The paper studies the interaction of users with search engines, which provides implicit feedback that can be used to extract useful knowledge and improve the quality of the search process. This feedback is encoded in the form of a query log that consists of a sequence of search actions, which contain information about submitted queries, documents viewed, and documents *clicked* by the users. In this paper, the authors propose characterizing documents and queries via the information available within a query log, with the goal of detecting either query polysemy or spam-hosts and spam-queries, i.e., queries that exhibit the undesirable property of showing a higher rate of spam pages in their list of results than other queries.

The last paper, “*Exploring the Impact of Profile Injection Attacks in Social Tagging Systems*” by Maryam Ramezani, J.J. Sandvig, Runa Bhaumik, Robin Burke, and Bamshad Mobasher [5], addresses the problem of vulnerability of social tagging systems to spamming and profile injection attacks as in the case of all open and adaptive systems that rely on user input to organize and present content. Malicious users may try to distort the system’s behavior by inserting erroneous or misleading annotations, thus altering the way in which information is presented to legitimate users. In this paper, the authors present and study two types of potential attacks against tagging systems. Using real data from a popular social tagging Web site, they empirically evaluate the impact of these attacks and their variants. Specifically, they consider two variants of an attack (called the overload attack) designed to promote a resource by adding different types of annotations to that resource, and another type of attack (called a piggyback attack) designed to promote a resource by associating it with other resources. Their results

show that current systems are vulnerable to attacks especially in scenarios where the attacker focuses on a specific target group of users to promote a target resource.

4.2 Web Personalization and User Profiling

The first paper, “*Detecting concept drift in web usage mining*” by Jose Borges and Mark Levene [6], studies the problem of detecting concept drift and assessing how much log data should be used to build a prediction model in the context of web usage mining, when predicting the user’s next web page access. To tackle this problem they use a variable length Markov chain to represent a collection of users’ web navigation sessions and to provide the choice of predictions for unseen sessions. The window size used to train the Markov chain model is a parameter that determines the amount of log data that is used to build the model. To evaluate the prediction accuracy they use the mean absolute error and the hit and miss scores. Intuitively, if the prediction accuracy gets better when training on a window of a particular size, rather than on the whole navigation history, then they have detected concept drift. Extensive experimental evaluation on real data sets support the conclusion that a change in users’ navigation behavior, i.e. concept drift, can be detected with the aid of the scoring metrics used to evaluate precision accuracy.

The second paper, winner of the Yahoo! Research Best Paper Award, “*Learning Preferences of New Users in Recommender Systems: An Information Theoretic Approach*” by Al M. Rashid, George Karypis, and John Riedl [7], addresses the challenge that a new user poses to Collaborative Filtering recommenders. Since the system has no knowledge about the preferences of the new user, it cannot provide personalized recommendations. A new user preference elicitation strategy needs to ensure that the user does not abandon a lengthy signup process, nor lose interest in returning to the site due to the low quality of initial recommendations. To solve this problem, they develop a set of information theoretic strategies for the new user problem. They propose an offline simulation framework, and evaluate the strategies through extensive offline simulations and an online experiment with real users of a live recommender system.

The third paper, “*Justified Recommendations based on Content and Rating Data*” by Panagiotis Symeonidis, Alex Nanopoulos, and Yannis Manolopoulos [8], addresses the limitations of existing recommendation justification systems which are based solely on navigational information, by proposing a novel approach that attains simultaneously accurate and justifiable recommendations. They construct a feature profile for the users, to reveal their favorite features. Moreover, they create biclusters (i.e. group of users which exhibit highly correlated ratings on groups of items) to exploit partial matching between the preferences of the test user and each community of users. They evaluate the quality of their justifications with an objective metric in a real data set, showing the superiority of the proposed approach. They also conduct a user study to compare users’ satisfaction using

the existing versus the proposed justification style. The study showed that the proposed justification style is the users’ favorite choice.

The fourth paper, “*Personalized Web Search by Using Learned User Profiles in Re-ranking*” by Jia Hu and Philip K Chan [9], addresses the challenge of distinguishing the different contexts of the same query due to individual users having different interests. To improve the relevance of search results, they propose re-ranking results based on a learned user profile and extending their previous work which introduced a scoring function for re-ranking search results based on a learned User Interest Hierarchy (UIH), and where the results indicated that they could improve relevance at lower ranks, but not at the top 5 ranks. In this paper, they improve the scoring function by incorporating new term characteristics, image characteristics, and pivoted length normalization.

4.3 Best paper award

WebKDD 2008 featured a best paper award, made possible with the kind support of Yahoo! Research. The paper was chosen by the workshop organizers among those that have been unanimously considered acceptable by all reviewers and received the best scores and comments. The Best Paper Award went to “*Learning Preferences of New Users in Recommender Systems: An Information Theoretic Approach*” by Al M. Rashid, George Karypis, and John Riedl. The paper focuses on recommender systems, one of the oldest and most successful application domains of Web usage mining. The authors take a new perspective (that of information theory) in dealing with the cold-start problem, i.e. the problem of delivering proper recommendations to users for which very little is known. This is a known and challenging problem, to which much effort has been devoted, yet, even in the 10th year of WebKDD there were not many satisfactory solutions to it. The best paper award, consisted of a winner certificate from by Yahoo! Research and \$500, and it was presented to the winners by Yahoo researcher, Aristidis Gionis.

4.4 Closing discussion

A prominent issue brought forward in the closing discussion of WebKDD 2008 was the need for datasets that are made publicly available and are appropriate for analysis, evaluation and comparisons. The need for public datasets has become increasingly pressing in the last years, and more so since many conferences in the area (including KDD) expect that the evaluation of methods can be done in a transparent and reproducible way. This need is even more acute in WebKDD research, because of the complexity of Web server logs and application logs, the absence of a straightforward ground truth in the data and the sophisticated nature of Web applications (including applications in the Social Web). It was also suggested to define competitions around published datasets, since this brings forward new ideas and new solutions.

It was stressed that knowledge discovery on the Web has become mainstream, and that sophisticated applications are proliferating, giving rise to complex new challenges, many of which are interdisciplinary in nature. In fact, “Web science”

as a new concept and as a conference series is an indication of the need for bringing together researchers that study the Web from different perspectives. The interaction with scholars from other disciplines is therefore seen as a necessity and as an opportunity: Interaction and understanding of the methodologies and solutions offered in each discipline is the basis for new ideas and innovative solutions.

The discussions concluded that WebKDD should sharpen its focus on modeling user behavior on the Web. As everyone noted, Web Mining in general had become a fairly mainstream topic. However, there is room within this area for a workshop to focus specifically on the modeling and the analysis of increasingly more complex Web environment and data sets. This makes the need for more sophisticated approaches for modeling or predicting user behavior even more pronounced. Indeed, this being the original focus of WebKDD, to some extent this would be a return to the roots, but in a much more complex WWW.

We also discussed creating and maintaining a single Web site associated with WebKDD on which, among other things, people can post an up-to-date, detailed, and annotated list of current research problems. In the long run, not only this will benefit the WebKDD research community, but good solutions to some of these problems can be the basis of a format for future workshops (or even a contest).

5. CONCLUSIONS AND FUTURE DIRECTIONS

The WebKDD workshop series has accompanied the new area of knowledge discovery on the Web from its early time to the present day. Mining the Web has now become mainstream, many of the early problems have been solved but many still remain open. Intelligent applications that exploit content, social applications that interconnect people and foster exchange, cooperation and flow of influence, high expectations in the retrieval of answers – rather than content, all these applications have emerged and brought new challenges to WebKDD. One of them is in protecting the Web and its users from malice – malicious content, malevolent services and their creators. A further one is privacy: The Social Web has become a great venue for interaction among people; people join it to share with others – via information, experiences, ideas, and feelings. As part of what was formerly private is now shared among friends, the notion of private “sphere” needs redefinition, but privacy protection will remain as indispensable as before.

An enumeration of grand challenges is beyond the scope of this report. As people join the Web and invent more and more applications in it, challenges emerge quite naturally. Interdisciplinary efforts are a key to understanding these challenges and a key to solving them. WebKDD has contributed to understanding and analyzing the Web for 10 years. We hope that the series will keep inspiring scholars and, most importantly, that the scholars who constitute the Web Mining Community will sustain it and continue contributing to understanding the Web and learning from it.

6. ACKNOWLEDGMENTS

We would like to thank the authors of all submitted papers whose creative work has made the competitive programme of

WebKDD 2008 possible. We are grateful to the PC members for their thorough and dedicated work: *Ricardo Baeza-Yates, Bettina Berendt, Indrajit Bhattacharya, Ed Chi, Wei Fan, Michael Hachslar, Xin Jin, George Karypis, Mark Last, Huan Liu, Ee-Peng Lim, Alex Nanopoulos, Rajeev Motwani, Nitendra Rajput, Tamas Sarlos, Fabrizio Silvestri, Andrew Tomkins*. Finally, we are grateful to Yahoo! Research for sponsoring the best paper award.

In the last 10 years, the WebKDD workshops were organized by several people, not all of whom were involved this year. Yet, they have shaped WebKDD to no less extent, and we are grateful to them for the hard and dedicated work that made 10 years of WebKDD possible, these include: *Brij Masand* who has been the first WebKDD workshop co-organizer, and had the idea of a workshop on Web usage mining at KDD. He deserves the gratefulness of all other organizers. We also acknowledge the rest of the past organizers: *Ronny Kohavi, Bing Liu, and Philip S. Yu*. We are indebted to the Workshop Chairs of the KDD conferences since 1999 for entrusting us with the organization of WebKDD, providing advice and active coordination support from the first announcement to the concerted action of preparing the workshop notes for local distribution. Springer-Verlag has been our reliable publisher for most of the workshops, entrusting us a volume in the prestigious LNCS series ever since the first, still unknown WebKDD workshop in 1999; and later, when WebKDD has reached maturity and regularity, launching the “Advances in Web Mining and Web Usage Analysis” subseries to accommodate the extended versions of the papers.

But most of all we are deeply indebted to the Web Mining Community, who has accompanied WebKDD for these 10 years, who has contributed with their submissions, reviews, ideas, comments and trust and made the first 10 years of knowledge discovery on the Web possible.

7. REFERENCES

- [1] Akshay Java, Anupam Joshi, and Tim Finin, “*Detecting Communities via Simultaneous Clustering of Graphs and Folksonomies*”, In Proceedings of WebKDD 2008, 2008 KDD Workshop on Web Mining and Web Usage Analysis, Las Vegas, NV, Aug. 2008.
- [2] Prasanna K Desikan and Jaideep Srivastava, “*I/O efficient computation of First Order Markov Measures for Large and Evolving Graphs*”, In Proceedings of WebKDD 2008, 2008 KDD Workshop on Web Mining and Web Usage Analysis, Las Vegas, NV, Aug. 2008.
- [3] Yiling Chen, David Pennock, and Tejaswi Kasturi, “*An Empirical Study of Dynamic Pari-mutuel Markets: Evidence from the Tech Buzz Game*”, In Proceedings of WebKDD 2008, 2008 KDD Workshop on Web Mining and Web Usage Analysis, Las Vegas, NV, Aug. 2008.
- [4] Carlos Castillo, Claudio Corsi, Debora Donato, Paolo Ferragina, and Aristides Gionis, “*Query-log mining for detecting polysemy and spam*”, In Proceedings of WebKDD 2008, 2008 KDD Workshop on Web Mining and Web Usage Analysis, Las Vegas, NV, Aug. 2008.
- [5] Maryam Ramezani, J.J. Sandvig, Runa Bhaumik, Robin Burke, and Bamshad Mobasher, “*Exploring the Impact of Profile Injection Attacks in Social Tagging Systems*”, In Proceedings of WebKDD 2008, 2008 KDD Workshop on Web Mining and Web Usage Analysis, Las Vegas, NV, Aug. 2008.

- [6] Jose Borges and Mark Levene, “*Detecting concept drift in web usage mining*”, In Proceedings of WebKDD 2008, 2008 KDD Workshop on Web Mining and Web Usage Analysis, Las Vegas, NV, Aug. 2008.
- [7] M. Rashid, George Karypis, and John Riedl, “*Learning Preferences of New Users in Recommender Systems: An Information Theoretic Approach*”, In Proceedings of WebKDD 2008, 2008 KDD Workshop on Web Mining and Web Usage Analysis, Las Vegas, NV, Aug. 2008.
- [8] Symeonidis, Alex Nanopoulos, and Yannis Manolopoulos, “*Justified Recommendations based on Content and Rating Data*”, In Proceedings of WebKDD 2008, 2008 KDD Workshop on Web Mining and Web Usage Analysis, Las Vegas, NV, Aug. 2008.
- [9] Jia Hu and Philip K Chan, “*Personalized Web Search by Using Learned User Profiles in Re-ranking*”, In Proceedings of WebKDD 2008, 2008 KDD Workshop on Web Mining and Web Usage Analysis, Las Vegas, NV, Aug. 2008.
- [10] B. Sarwar, G. Karypis, J. Konstan, “*Application of Dimensionality Reduction in Recommender System - A Case Study*”, In Proceedings of WebKDD 2000, 2000 KDD Workshop on Web Mining for e-commerce-Challenges and Opportunities, Boston, MA, 2000.
- [11] Jose Borges and Mark Levene, “*Data Mining of User Navigation Patterns*”, In Proceedings of WebKDD 1999, 1999 KDD Workshop on Web Usage Analysis and User Profiling, San Diego, CA, 1999.
- [12] L. Becchetti C. Castillo D. Donato S. Leonardi R. Baeza-Yates, “*Using Rank Propagation and Probabilistic Counting for Link-Based Spam Detection*”, In Proceedings of WebKDD 2006, 2006 KDD Workshop on Web Usage Analysis and User Profiling, Philadelphia, PA, 2006.

About the authors:

Olfa Nasraoui is the Endowed Chair of E-commerce and the Director of the Knowledge Discovery and Web Mining Lab at the University of Louisville, where she is also Associate professor in Computer Science and Engineering. She received her Ph.D. in Computer Engineering and Computer Science from the University of Missouri-Columbia in 1999. From 2000 to 2005, she was an Assistant Professor at the University of Memphis. Her research activities include Data Mining, in particular, Web mining and mining evolving data streams, Personalization, and Computational Intelligence. She has served on the organizing and program committees of several conferences and workshops, including WebKDD 2004-2008. She is the recipient of the National Science Foundation CAREER Award for outstanding young scientists. Her research is funded mainly by NSF and by NASA. She is a member of IEEE, IEEE WIE, and ACM. (<http://www.louisville.edu/~o0nasr01>)

Myra Spiliopoulou is Professor at the Faculty of Computer Science, Otto-von-Guericke-University Magdeburg, Germany. Her research is on the development and enhancement of knowledge discovery methods for person-web interaction, for text analysis and for knowledge management. Her research on web usage mining and data preparation, text mining, temporal mining and pattern evolution has appeared in journals, conferences, books and workshops. She is regular reviewer in many data mining conferences, including KDD, ECML/PKDD and SIAM Data Mining, of the IEEE TKDE

Journal and of many workshops in subjects associated to web mining. She has co-organized most of the WEBKDD workshops in the KDD conference series since 1999. In the ECML/PKDD conference series she has also given several tutorials on web mining since 1999. In March 2005, she organized the 29th Annual Conference of the German Classification Society (GfKI'2005) in Magdeburg, and in 2006 she was PC Co-Chair of the ECML/PKDD that took place in Berlin (Sept. 2006) and PC Co-Chair of this year's NLDB conference on natural language processing and databases, which took place in London (June 2008). (<http://omen.cs.uni-magdeburg.de/itikmd>)

Osmar R. Zaiane is a McCalla Professor at the Department of Computing Science, University of Alberta, Canada, and director of the Alberta Ingenuity Centre for Machine Learning (AICML). Zaiane earned his Ph.D. in Computing Science at Simon Fraser University in 1999 under the supervision of Dr. Jiawei Han focusing on Web Mining and Multimedia Mining. He serves on the editorial boards of several prominent journals, including Knowledge and Information Systems, An International Journal. He is also the Editor-in-Chief of ACM SIGKDD Explorations. He has research interests in novel data mining algorithms, web mining, text mining, image mining, and information retrieval. His current centre of attention is health informatics and the application of data mining in health care and medical research. He has published more than 100 papers in refereed international conferences and journals, and taught on all six continents. Zaiane was the program co-chair of the IEEE International Conference on Data Mining. He co-chaired the ACM SIGKDD WebKDD workshops four times in the last 10 years (2002, 2003, 2005 and 2008). He has been an ACM Member since 1986. (<http://www.cs.ualberta.ca/~zaiane/>)

Jaideep Srivastava is a Professor of Computer Science & Engineering at the University of Minnesota. He has established and led a laboratory that has conducted research in databases, multimedia systems, and data mining. He has supervised 23 Ph.D. dissertations and 44 MS theses, and has authored/co-authored over 185 papers in journals and conferences. Dr. Srivastava has an active collaboration with the technology industry, both for research and technology transfer, and is an often-invited participant in technical and technology strategy forums. Dr. Srivastava's industry experience includes data mining at Amazon.com and at Yodlee. He has provided technology and technology strategy advice to a number of large corporations, including Cargill, United Technologies, IBM, Honeywell, 3M, and Persistent Systems. He is on the editorial board of a number of journals, and has served as Program and Conference Chair for a number of prominent conferences. Dr. Srivastava has recently been appointed as the Senior Technology Advisor for the State of Minnesota, where he provides advice on information technologies to the State Chief Information Officer. He has been elected a Fellow of the IEEE, and has been appointed a Distinguished Visitor by the IEEE Computer Society. He has a Ph.D. from the University of California, Berkeley, and bachelors from IIT Kanpur, India. (<http://www.cs.umn.edu/people/faculty.php?user=srivasta>)

Bamshad Mobasher is a Professor of Computer Science and the director of the Center for Web Intelligence at the School of Computing of DePaul University in Chicago. His research areas include Web mining, Web personalization, predictive user modeling, agent-based systems, and information retrieval. He has published more than 100 scientific articles, numerous book chapters, and several edited books in these areas. Dr. Mobasher received his PhD at Iowa State University in 1994. Prior to DePaul he was an assistant professor of Computer Science at the University of Minnesota, Twin Cities, where he did some seminal work on Web mining and started some of the first research groups focusing on *Web Usage Mining*. As the director of the Center for Web Intelligence, Dr. Mobasher is directing research in Web mining, Web analytics, and personalization, as well as

overseeing several related joint projects with the industry. His most recent activities include an edited volume, "Intelligent Techniques for Web Personalization", published by Springer, culminating from a series of successful workshops at IJCAI and AAI on the same topic; and a special issue of *ACM Transactions on Internet Technologies* on Web personalization. He has also been one of the organizers of the WebKDD workshops on Knowledge discovery on the Web which have been held at ACM SIGKDD conference for the past 10 years. He is currently editing a special issue of *User Modeling and User Adapted Interaction* on Data Mining for Personalization, and he is the program co-chair of the 2008 ACM International Conference on Recommender Systems.

(<http://maya.cs.depaul.edu/~mobasher>)