Web Communities

Yanchun Zhang · Jeffrey Xu Yu · Jingyu Hou

Web Communities

Analysis and Construction

With 28 Figures

Springer

Authors
Yanchun Zhang

School of Computer Science and Mathematics
Victoria University of Technology
Ballarat Road, Footscray
PO Box 14428
MC 8001, Melbourne City, Australia
yzhang@csm.vu.edu.au

Jeffrey Xu Yu

Dept. of Systems Engineering and Engineering Management
Chinese University of Hong Kong
Shatin, N.T., Hong Kong, China
yu@se.cuhk.edu.hk

Jingyu Hou

School of Information Technology
Deakin University
Burwood, Victoria 3125, Australia
jingyu@deakin.edu.au

Library of Congress Control Number: 2005936102

ACM Computing Classification (1998): H.3, H.5

ISBN-10 3-540-27737-4 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-27737-8 Springer Berlin Heidelberg New York

Typeset by the authors using a Springer TEX macro package
Production: LE-TEX Jelonek, Schmidt & Vöckler GbR, Leipzig
Cover design: KünkelLopka Werbeagentur, Heidelberg

Printed on acid-free paper 45/3142/YL - 5 4 3 2 1 0

Dedication

To Jinli and Dana
From Yanchun

To Hannah, Michael and Stephen
From Jeffrey

To Huiming, Mingxi and Mingyi
From Jingyu

Contents

Preface

The rapid development of Web technology has made the World Wide Web an important and popular application platform for disseminating and searching information as well as conducting business. However, due to the lack of uniform schema for Web documents, the low precision of most search engines and the information explosion on the World Wide Web, the user is often flooded with a huge amount of information.

Unlike the conventional database management in which data models and schemas are defined, the Web community, which is a set of Web-based objects (documents and users) that has its own logical structures, is another effective and efficient approach to reorganize Web-based objects, support information retrieval and implement various applications. According to the practical requirements and concerned situations, the Web community would appear as different formats.

This book addresses the construction and analysis of various Web communities based on information available from Web, such as Web document content, hyperlinks, semantics and user access logs. Web community applications are another aspect emphasized in this book. Before presenting various algorithms, some preliminaries are provided for better understanding of the materials. Representative algorithms for constructing and analysing various Web communities are then presented and discussed. These algorithms, as well as their discussions, lead to various applications that are also presented in this book. Finally, this book summarizes the main work in Web community research and discusses future research in this area.

Acknowledgements

Our special thanks go to Mr. Guandong Xu and Mr. Yanan Hao for their assistance in preparing manuscripts of the book.