Contents

i. Workshop Committee
ii. Introduction

Usage Analysis

A Large-scale Study of Automated Web Search Traffic ................................................. 1
Greg Buehrer, Jack Stokes and Kumar Chellapilla

Identifying Web Spam with User Behavior Analysis ....................................................... 9
Yiqun Liu, Rongwei Cen, Min Zhang, Liyun Ru and Shaoping Ma

Query-log mining for detecting spam ........................................................................ 17
Carlos Castillo, Claudio Corsi, Debora Donato, Paolo Ferragina and Aristides Gionis

Text Analysis

Cleaning Search Results using Term Distance Features ................................................. 21
Josh Attenberg and Torsten Suel

Exploring Linguistic Features for Web Spam Detection: A Preliminary Study ................ 25
Jakub Piskorski, Marcin Sydow and Dawid Weiss

Latent Dirichlet Allocation in Web Spam Filtering ......................................................... 29
István Bíró, Jácint Szabó and András A. Benczúr

General

Analysing Features of Japanese Splogs and Characteristics of Keywords ...................... 33
Yuuki Sato, Takehito Utsuro, Tomohiro Fukuhara, Yasuhide Kawada, Yoshiaki Murakami, Hiroshi Nakagawa and Noriko Kando

Webspam Identification Through Content and Hyperlinks ........................................ 41
Jacob Abernethy, Olivier Chapelle and Carlos Castillo

Social Networks

Identifying Video Spammers in Online Social Networks ............................................... 45
Fabricio Benevenuto, Tiago Rodrigues, Virgilio Almeida, Jussara Almeida, Chao Zhang and Keith Ross

A Few Bad Votes Too Many? Towards Robust Ranking in Social Media ...................... 53
Jiang Bian, Yandong Liu, Eugene Agichtein and Hongyuan Zha

The Anti-Social Tagger – Detecting Spam in Social Bookmarking Systems ................ 61
Beate Krause, Christoph Schmitz, Andreas Hotho and Gerd Stumme

Link Analysis

Robust PageRank and Locally Computable Spam Detection Features ............................ 69
Reid Andersen, Christian Borgs, Jennifer Chayes, John Hopcroft, Kamal Jain, Vahab Mirrokni, and Shang-Hua Teng