

*Commenced Publication in 1973*

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

## Editorial Board

David Hutchison

*Lancaster University, UK*

Takeo Kanade

*Carnegie Mellon University, Pittsburgh, PA, USA*

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

*Massachusetts Institute of Technology, MA, USA*

Demetri Terzopoulos

*New York University, NY, 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*

Henry S. Baird Daniel P. Lopresti (Eds.)

# Human Interactive Proofs

Second International Workshop, HIP 2005  
Bethlehem, PA, USA, May 19-20, 2005  
Proceedings

## Volume Editors

Henry S. Baird  
Daniel P. Lopresti  
Lehigh University  
Computer Science and Engineering Department  
Bethlehem, Pennsylvania, USA  
E-mail: {baird,lopresti}@cse.lehigh.edu

Library of Congress Control Number: 2005926089

CR Subject Classification (1998): I.4, I.5, H.4, C.2, D.4.6, K.4.4, K.6.5, H.3

|         |   |
|---------|---|
| ISSN    | 0302-9743   |
| ISBN-10 | 3-540-26001-3 Springer Berlin Heidelberg New York     |
| ISBN-13 | 978-3-540-26001-1 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  
[springeronline.com](http://springeronline.com)

© Springer-Verlag Berlin Heidelberg 2005  
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Boller Mediendesign  
Printed on acid-free paper      SPIN: 11427896      06/3142      5 4 3 2 1 0

# Preface

E-commerce services are suffering abuse by programs (bots, spiders, etc.) masquerading as legitimate human users. Efforts to defend against such attacks have, over the past several years, stimulated investigations into a new family of security protocols – “Human Interactive Proofs” (HIPs) – which allow a person to authenticate herself as a member of a given group: e.g., as a human (vs. a machine), as herself (vs. anyone else), as an adult (vs. a child). Most commercial uses of HIPs today are CAPTCHAs, “Completely Automatic Public Turing tests to tell Computers and Humans Apart,” which exploit the gap in ability between humans and machine vision systems in reading images of text. HIP challenges can also be non-graphical, e.g., requiring recognition of speech, solving puzzles, etc.

We are pleased to present the first refereed and archivally published collection of state-of-the-art papers on HIPs and CAPTCHAs. Each paper was reviewed by three members of the Program Committee, judged by the Co-chairs to be of sufficient relevance and quality, and revised by the authors in response to the referees’ suggestions.

The papers investigate performance analysis of novel CAPTCHAs, HIP architectures, and the role of HIPs within security systems. Kumar Chellapilla, Kevin Larson, Patrice Simard, and Mary Czerwinski describe user trials of a CAPTCHA designed to resist segmentation attacks, including a systematic evaluation of its tolerance by human users. Henry Baird, Michael Moll, and Sui-Yu Wang analyze data from a human legibility trial of another segmentation-resistant CAPTCHA and locate a highly legible engineering regime. Amalia Rusu and Venu Govindaraju describe research towards CAPTCHAs based on reading synthetically damaged images of real images of unconstrained handwritten text. Yong Rui, Zicheng Liu, Shannon Kallin, Gavin Janke, and Cem Paya discuss the results of experiments with human subjects presented with two kinds of CAPTCHAs: one based on reading text, and a new one based on the detection of well-formed synthetic faces.

Monica Chew and J.D. Tygar discuss collaborative filtering CAPTCHAs which do not depend on absolute answers, but are graded by comparison with other people’s answers. Tim Converse proposes CAPTCHA generation as a not-for-profit Web service and argues for open-sourcing the code. Daniel Lopresti proposes using instances of open pattern recognition problems to build CAPTCHAs in order to benefit both online security and pattern recognition research.

Jon Bentley and Colin Mallows describe methods for quantifying the assurance that can be inferred from a correct answer to a password query: the principles underlying this analysis are applicable to the evaluation of CAPTCHA security. Rachna Dhamija and J.D. Tygar investigate HIPs in which a user issues challenges to the computer, rather than the other way around, enabling the detection of phishing attacks.

We are warmly grateful for the time and skill volunteered by our Program Committee, as well as to our Advisory Board for kindly assisting in publicizing the event and suggesting ways to make the program more stimulating.

May 2005

Henry S. Baird  
Daniel P. Lopresti  
Workshop Co-chairs, HIP 2005  
Department of Computer Science & Engineering  
Lehigh University  
Bethlehem, PA, USA

# Organization

HIP 2005 was organized by the Department of Computer Science & Engineering, Lehigh University and was endorsed by IAPR, the International Association for Pattern Recognition.

## Advisory Board

|                |  |
|----------------|--|
| Jon Bentley    | Avaya Labs Research                          |
| Manuel Blum    | Computer Science, Carnegie-Mellon University |
| Andrei Broder  | IBM Research                                 |
| Gordon Legge   | Psychology Depts., University of Minnesota   |
| Richard Lipton | College of Computing, Georgia Tech           |
| Patrice Simard | Microsoft Research                           |
| Doug Tygar     | Computer Science & SIMS Depts., UC Berkeley  |

## Program Committee

|                       |  |
|-----------------------|--|
| Luis von Ahn          | Computer Science, Carnegie-Mellon University |
| Kumar Chellapilla     | Microsoft Research                           |
| Monica Chew           | Computer Science, UC Berkeley                |
| James Clark           | ECE Dept, McGill University                  |
| Brian Davison         | Computer Science & Engineering, Lehigh       |
| Irfan Essa            | GVU Center, Georgia Tech                     |
| Venugopal Govindaraju | CEDAR, SUNY Buffalo                          |
| Greg Kochanski        | Phonetics Lab, Oxford University             |
| Richard Landsman      | America Online                               |
| Cem Paya              | Microsoft MSN/Passport                       |
| David Pletcher        | Lawrence Livermore Lab, Berkeley             |
| Chilin Shih           | EALC & Linguistics, University of Illinois   |
| Richard Sproat        | Linguistics & ECE, University of Illinois    |

## Sponsoring Institutions

We gratefully acknowledge financial support from Microsoft Research and Avaya Labs Research.

# Table of Contents

## CAPTCHAs and Performance Analysis

|   |    |
|---|----|
| Building Segmentation Based Human-Friendly Human Interaction Proofs (HIPs) .....      | 1  |
| <i>K. Chellapilla, K. Larson, P.Y. Simard, and M. Czerwinski (Microsoft Research)</i> |    |
| A Highly Legible CAPTCHA That Resists Segmentation Attacks .....                      | 27 |
| <i>H.S. Baird, M.A. Moll, and S.-Y. Wang (Lehigh University)</i>                      |    |
| Visual CAPTCHA with Handwritten Image Analysis .....                                  | 42 |
| <i>A. Rusu and V. Govindaraju (University at Buffalo)</i>                             |    |
| Characters or Faces: A User Study on Ease of Use for HIPs .....                       | 53 |
| <i>Y. Rui, Z. Liu, S. Kallin, G. Janke, and C. Paya (Microsoft)</i>                   |    |

## HIP Architectures

|  |    |
|--|----|
| Collaborative Filtering CAPTCHAs .....                               | 66 |
| <i>M. Chew and J.D. Tygar (University of California at Berkeley)</i> |    |
| CAPTCHA Generation as a Web Service .....                            | 82 |
| <i>T. Converse (Yahoo!)</i>  |    |
| Leveraging the CAPTCHA Problem .....                                 | 97 |
| <i>D. Lopresti (Lehigh University)</i>                               |    |

## HIPs Within Security Systems

|   |     |
|---|-----|
| How Much Assurance Does a PIN Provide? .....                            | 111 |
| <i>J. Bentley and C. Mallows (Avaya Labs Research)</i>                  |     |
| Phish and HIPs: Human Interactive Proofs to Detect Phishing Attacks ..  | 127 |
| <i>R. Dhamija and J.D. Tygar (University of California at Berkeley)</i> |     |
| Author Index .....  | 143 |