Lecture Notes in Computer Science 7053

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

Pascal Bouvry
Mieczysław A. Kłopotek
Franck Leprévost
Małgorzata Marciniak
Agnieszka Mykowiecka
Henryk Rybiński (Eds.)

Security and Intelligent Information Systems

International Joint Conference, SIIS 2011
Warsaw, Poland, June 13-14, 2011
Revised Selected Papers

Springer

Volume Editors

Pascal Bouvry
ILIAS, University of Luxembourg
E-mail: pascal.bouvry@uni.lu

Mieczysław A. Kłopotek
Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland
E-mail: klopotek@ipipan.waw.pl

Franck Leprévost
University of Luxembourg
E-mail: franck.leprevost@uni.lu

Małgorzata Marciniak
Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland
E-mail: malgorzata.marciniak@ipipan.waw.pl

Agnieszka Mykowiecka
Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland
E-mail: agnieszka.mykowiecka@ipipan.waw.pl

Henryk Rybiński
Institute of Computer Science, Warsaw University of Technology, Poland
E-mail: h.rybinski@ii.pw.edu.pl

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-25260-0 e-ISBN 978-3-642-25261-7
DOI 10.1007/978-3-642-25261-7
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011944280

CR Subject Classification (1998): C.2.4, C.2.5, H.2.8, I.2.9, K.3.1, K.4.4, I.5

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

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This volume contains papers selected from those accepted for presentation at the International Joint Conference on Security and Intelligent Information Systems (SIIS) which was held in Warsaw, Poland, June 13–14, 2011. The conference was organized by the Institute of Computer Science of the Polish Academy of Sciences, the Institute of Computer Science at the Warsaw University of Technology, and the University of Luxembourg as a joint meeting of the Second Luxembourg-Polish Meeting on Security and Trust and the 19th International Conference on Intelligent Information Systems.

The papers submitted to the conference were organized into three thematic tracks: security and trust, data mining and machine learning, and natural language processing. Based on anonymous peer-review of the 60 submissions, 29 papers were accepted for inclusion in this volume. Moreover, the volume contains two invited papers by Gerhard Frey and Joakim Nivre.

New technologies have emerged in these last few decades. They allow better and easier collaboration and interaction between people, which in turn leads to new threats and dangers. In order to enable any economic or social exchange, trust between the actors is a necessity. A formal definition of trust is needed in this context, and the means to enable and measure trust are required. It is also necessary to fight all potential crimes and attacks. Security and trust are therefore two very key topics in our society. Under these circumstances, it is no wonder that the **Security and Trust** track of the SIIS 2011 joint conference was given a *primus inter pares* status.

The scientific part of the SIIS 2011 conference was opened by an invited plenary talk given by Gerhard Frey entitled "Is Arithmetic Geometry Necessary for Public-Key Cryptography?". The talk presented the challenges of current public-key cryptology based discrete logarithm problems (DLP) in finite cyclic groups. G. Frey and E. Kani's paper entitled "Correspondences on Hyperelliptic Curves and Applications to the Discrete Logarithm Problem" — published in these proceedings — addresses these issues. More precisely, the authors recall that divisor class groups of carefully chosen curves over finite fields provide the main source of groups for DLP. They also recall that curves of genus $g \geq 4$ and non-hyperelliptic curves of genus $g = 3$ have to be avoided for security reasons. Furthermore, Smith showed that 'many' hyperelliptic curves of genus 3 have to be avoided too. The deep reason is due to the existence of isogenies of low degrees between the Jacobians of these hyperelliptic curves to the Jacobians of non-hyperelliptic curves (of the same genus), and hence the DLP is 'easily' transferred from one Jacobian to another. G. Frey and E. Kani take the point of view of correspondences and isogenies: for each g, their paper describes how to find a Hurwitz space parametrizing a subspace of those hyperelliptic curves C of genus g which admit a non-trivial correspondence to a curve D of genus g that

can be expected to be non-hyperelliptic. Their approach is purely geometric at the beginning (where they assume the ground field to be algebraically closed) and they focus on rationality issues in a second step. In the frontier case $g = 3$, they give a parametrization of these hyperelliptic curves of genus 3 in terms of a Hurwitz moduli space with monodromy group S_4, the symmetric group on 4 letters. In particular, they recover Smith's results, and announce that a future paper of E. Kani extends these results to the situation where the ground field has the characteristic 2. G. Frey and E. Kani's important paper, together with other results, leads to the conclusion stated at the end of his talk: according to today's knowledge, it is safer to avoid curves of genus ≥ 3 for cryptographic purposes and use only elliptic curves or (simple) Jacobians of genus 2 curves.

The **Security and Trust** Track, opened by G. Frey's talk, also included presentations of the following eight papers. In terms of trust metrics in modern ad-hoc networks, M. Seredyński et al. present in "Solving Soft Security Issues in MANETs Using an Evolutionary Approach" new approaches where decentralized strategies are discussed for trust computation. The paper "Camera Sabotage Detection for Surveillance Systems" by D. Ellwart et al. describes new methods for detecting anomalies in camera surveillance systems. A. Poniszewska-Maranda, in the paper "Implementation of Access Control Model for Distributed Information Systems Using Usage Control," presents an enhanced model for access control based on extensions of role-based access control that enables more dynamic management. Nowadays, cryptography is able to help in authentication, signature, encryption, and non-repudiation. But additional mechanisms are required to help erase all available traces of information exchange between partners. The paper "Beyond TOR: The TrueNyms Protocol" by N. Bernard and F. Leprévost introduces TrueNyms that allows the masking of all information during the exchange on an encoded channel that might remain in the packet headers (e.g., source of the packet, number of packets, etc.). The paper "A Signature Scheme for Distributed Executions Based on Control Flow Analysis" by S. Varrette et al. describes a way to help certify the results on distributed platforms such as desktop-based grids. In the paper "Computational Aspects of Attack-Defense Trees," B. Kordy et al., after introducing the extension of attack-trees called attack-defense trees, demonstrate that the computational complexity of this approach remains equivalent to those of attack trees. D. Priemuth-Schmid presents two attacks on simplifed versions of the stream cipher K2 which was introduced at SECRYPT 2007 by S. Kiyomoto, T. Tanaka, and K. Sakurai.

Data Mining and Machine Learning (DM and ML) was the next conference track. It was opened by an invited plenary talk by Alessio Lomuscio entitled "Verification of Multi-Agent Systems." Serial and parallel algorithms for symbolic model checking for temporal-epistemic logic as well as bounded-model checking procedures were discussed in the talk. Moreover, applications of the methodology to the automatic verification of security protocols, Web services, and fault-tolerance were surveyed. (The paper elaborating on the issues presented in the talk will appear elsewhere and hence it is not included in this volume.)

Regarding DM and ML methods, several new methods of model discovery from data are presented. The paper "Model Selection in Logistic Regression Using p-Values and Greedy Search" by J. Mielniczuk and P. Teisseyre proposes a new method of model selection from a set of candidate models, and demonstrates its effectiveness and applicability in step-wise model construction. The paper "Landau Theory of Meta-Learning" by D. Plewczynski aims at creating and applying machine-learning algorithms in such a way that, for a given problem under scrutiny, a multitude of slightly different decision models can be derived, which can then make final decisions based on a majority vote. The author lists a number of such algorithms, proposes new ones, as well as their combinations, and demonstrates their good properties. In the paper "Multi-Test Decision Trees for Gene Expression Data Analysis," M. Czajkowski at al. argue that the voting should be performed at a single attribute level. It deals with ways to surpass the known problem with proper model construction when a number of different attributes have a similar predictive capability. Under such circumstances decision trees may perform worse than other approaches. The authors claim and demonstrate that one can keep the explanatory power of decision trees while at the same time making more reliable decisions by letting many tests vote at a given branching point of the tree. The paper "Rule-Based Approach to Computational Stylistics" by U. Stańczyk suggests on the other hand that one can start building a model (say, a classifier) filling it with a multitude of constituent rules of varying quality, and then identifying poorly performing features and removing rules which contain these features. Significant increases in decision quality are observed.

Another group of papers, pertaining to DM and ML tools, is devoted to issues in evolutionary optimization. The paper "Differential Evolution for High-Scale Dynamic Optimization" by M. Raciborski at al. proposes and explores a new area of application for differential evolution, showing its reliability in tasks with a dynamically changing environment. The paper "Towards an OpenCL Implementation of Genetic Algorithms on GPUs" by T. Puźniakowski and M. Bednarczyk deals with the technical side of the performance of genetic algorithms, demonstrating that the proper choice of an offspring selection method may provide a significant speed-up of the optimization process due to the technical properties of graphic cards. The paper "Evolutionary Algorithm Parameter Tuning with Sensitivity Analysis" by F. Pinel et al. tackles the delicate issue of tuning the many parameters of a typical evolutionary algorithm. The key idea is that sensitivity analysis allows us to identify the parameters that most strongly influence the performance for a given application, allowing the researcher to concentrate on tuning them properly.

The last group of papers in the DM and ML track is application-oriented. The paper "Image Recognition System for Diagnosis Support of Melanoma Skin Lesion" by W. Paja et al. deals with the application of image understanding in the medical domain, in particular for computer-aided automated classification of melanocytic skin lesions. Instead of a simple classification scheme, a mechanism for chaining diverse image processing methods is developed in order to

extract features from images. The paper "Playing in Unison in the Random Forest" by A. Wieczorkowska et al. addresses the issues related to the identification of instruments of an orchestra in the very difficult case of unison play (same tune for each instrument). It turns out that the techniques to be applied and features to be used differ significantly from those used for the recognition of a single instrument. Random forest classifiers are trained and used in the identification process. The paper "Scale Invariant Bipartite Graph Generative Model" by S. Chojnacki and M. Kłopotek is devoted to an important issue of modelling social networks with different modalities, such as user-item, author-paper, or actor-film networks. Traditionally used random graph models failed to represent some important aspects of such networks, such as node degree distributions in conjunction with clustering behavior. The newly introduced mechanisms allow for a much easier fitting of a model to real-world data. The last two papers address text/Web mining problems. The paper "Introducing Diversity to Log-Based Query Suggestions to Deal with Underspecified User Queries" by M. Sydow et al. explores the application of a concept of document set diversification, in order to improve responses to search engine queries. It turns out that to achieve a diversification in the response, there is no need to recall the original documents and one can rely on characteristics of previous queries only, enhanced possibly with some Wikipedia-based statistics. Wikipedia data are also valuable when categorizing documents, as K. Ciesielski et al demonstrate in the paper "Wikipedia-Based Document Categorization." A mapping between the words of a language and hierarchical Wikipedia categories is created on the basis of a Wikipedia category graph and page graph. It constitutes a foundation of mapping the whole document to a set of categories which is then rectified based on common supercategories and *tfidf* (term frequency inverse document frequency) like statistics.

The last track of the SIIS 2011 Joint Conference was devoted to **Natural Language Processing** (NLP). It was opened by an invited plenary talk by Joakim Nivre entitled "Bare-Bones Dependency Parsing." The author presented the general methodological and implementational issues connected with inducing parsers on the basis of annotated examples. In contrast to many other experiments, the presented approach does not utilize intermediate phrase structures.

In the Internet era, when more and more electronic texts in many natural languages become available each day, automatic processing of these texts is one of the most important tasks for computer applications. Among the two main approaches to NLP application building — rule-based and machine learning paradigms — the latter has become more popular. These methods are a common denominator of the first group of papers.

Regarding the problems of dependency parsing techniques, in addition to the invited paper by J. Nivre, A. Wróblewska and M. Woliński present preliminary experiments in the induction of a dependency parser for Polish. Although such experiments were already conducted for many languages, there were no results reported for Polish data yet. The next three papers in the NLP section concern various aspects of dealing with natural language semantics. In the first one

B. Broda et al. describe an evaluation methodology for automated Wordnet expansion algorithms. The next paper by Ł. Kobyliński addresses the problem of word sense disambiguation in a limited domain (in this case economy). The author uses class association rules to create an effective and human-understandable rule-based classifier. The third paper is devoted to semantic relations extraction. In this paper, A. Pohl describes an ontology-based method for selecting testing examples for relation extraction, and a method of their validation.

The subsequent four papers in the NLP section describe problems concerning words and phrases: M. Marcińczuk et al. describe the recognition of proper names in texts, using a very rich set of features for training CRF models. E. Hajnicz describes the creation of a semantic valence dictionary. T. Śniatowski and M. Piasecki present the outcomes of combining the results of three Polish taggers. Finally, the problems of lemmatization of nominal phrases in Polish are presented by Ł. Degórski.

In the final two papers included in the NLP section M. Junczys-Dowmunt and A. Szał present the concept of symmetrical word alignment, which outperforms one-way alignment, and A. Wawer and K. Sakwerda describe an experiment with building an ontology for sentiment analysis in the process of text annotation.

We would like to express our thanks to the invited speakers and the authors of papers for their contributions. We would also like to thank all the Program Committee members and invited reviewers for their excellent job. Last but not least, we gratefully acknowledge the generous support from the Office of Naval Research Global and Fonds National de la Recherche Luxembourg.

Pascal Bouvry
Mieczysław A. Kłopotek
Franck Leprévost
Małgorzata Marciniak
Agnieszka Mykowiecka
Henryk Rybiński

Conference Organization

Steering committee

Pascal Bouvry	University of Luxembourg
Mieczysław A. Kłopotek	Institute of Computer Science PAS, Poland
Jacek Koronacki	Institute of Computer Science PAS, Poland
Franck Leprévost	University of Luxembourg
Józef Lubacz	Warsaw University of Technology, Poland
Małgorzata Marciniak	Institute of Computer Science PAS, Poland
Mieczysław Muraszkiewicz	Warsaw University of Technology, Poland
Agnieszka Mykowiecka	Institute of Computer Science PAS, Poland
Björn Ottersten	University of Luxembourg
Henryk Rybiński	Warsaw University of Technology, Poland
Mirosław Słomiński	Warsaw University of Technology, Poland

Publishing Chair

Leonard Bolc	Polish Japanese Institute of Information Technology, Poland

Programme Committee

Witold Abramowicz	Poznań University of Economics, Poland
Stanisław Ambroszkiewicz	Institute of Computer Science PAS, Poland
Alex Biryukov	University of Luxembourg
António Horta Branco	University of Lisbon, Portugal
Luis Miguel de Campos	University of Granada, Spain
Andrzej Czyżewski	Gdańsk University of Technology, Poland
Jan Daciuk	Gdańsk University of Technology, Poland
Tapio Elomaa	Tampere University of Technology, Finland
Piotr Gawrysiak	Warsaw University of Technology, Poland
Marek Gorgoń	AGH University of Science and Technology, Poland
Jerzy W. Grzymała-Busse	University of Kansas, USA
Wojciech Jamroga	University of Luxembourg
Józef Korbicz	University of Zielona Góra, Poland
Zbigniew Kotulski	Warsaw University of Technology, Poland

Anna Kupść	University of Bordeaux 3, France
Alessio Lomuscio	Imperial College London, UK
Ramón López-Cózar Delgado	University of Granada, Spain
Krzysztof Marasek	Polish Japanese Institute of Information Technology, Poland
Nicola Di Mauro	University of Bari, Italy
Sjouke Mauw	University of Luxembourg
Maciej Michalewicz	IBM Netezza Poland
Zbigniew Michalewicz	University of Adelaide, Australia
Karel Pala	Masaryk University, Czech Republic
Wojciech Penczek	Institute of Computer Science PAS, Poland
Maciej Piasecki	Wrocław University of Technology, Poland
Gábor Prószéky	MorphoLogic, Hungary
Adam Przepiórkowski	Institute of Computer Science PAS, Poland
Zbigniew W. Raś	University of North Carolina at Charlotte, USA
Jan Rauch	University of Economics, Czech Republic
Gilbert Ritschard	University of Geneva, Switzerland
Peter Ryan	University of Luxembourg
Abdel-Badeeh M. Salem	Ain Shams University, Egypt
Franciszek Seredyński	Institute of Computer Science PAS, Poland
Andrzej Skowron	Warsaw University, Poland
Marian Srebrny	Institute of Computer Science PAS, Poland
Marcin Sydow	Polish Japanese Institute of Information Technology, Poland
Stan Szpakowicz	University of Ottawa, Canada
Roman Świniarski	San Diego State University, USA
Ryszard Tadeusiewicz	AGH University of Science and Technology, Poland
Jonathan Timmis	University of York, UK
Krzysztof Trojanowski	Institute of Computer Science PAS, Poland
Angelina Tzacheva	University of South Carolina Upstate, USA
Jerzy Urbanowicz	Institute of Computer Science PAS, Poland
Antti Valmari	Tampere University of Technology, Finland
Zygmunt Vetulani	Adam Mickiewicz University, Poland
Alicja Wakulicz-Deja	University of Silesia, Poland
Sławomir T. Wierzchoń	Institute of Computer Science PAS, Poland
Peter Wittenburg	Max Planck Institute for Psycholinguistics, The Netherlands
Karsten Wolf	University of Rostock, Germany
Bożena Woźna-Szcześniak	Jan Długosz University, Poland

Invited Reviewers

Nicolas Bernard
Elżbieta Hajnicz
Eric Joanis
Andrew V. Jones
Hugo Jonker
Alistair Kennedy
Mirosław Kurkowski
Jacek Małyszko
Marek Ostaszewski
Jakub Piskorski
Peter Ryan
Marcin Seredyński
Jarosław Skaruz
Mirosław Szaban
Piotr Świtalski
Sebastien Varrette
Alina Wróblewska

Organizing Committee

Piotr Borkowski
Michał Ciesiołka
Grzegorz Mańko
Marek Miszewski
Antoni Siennicki

Table of Contents

Invited Papers

Security and Trust

Data Mining and Machine Learning

Natural Language Processing