Lecture Notes in Networks and Systems 532

Pablo García Bringas · Hilde Pérez García ·
Francisco Javier Martínez de Pisón ·
José Ramón Villar Flecha · Alicia Troncoso Lora ·
Enrique A. de la Cal · Álvaro Herrero ·
Francisco Martínez Álvarez · Giuseppe Psaila ·
Héctor Quintián · Emilio Corchado *Editors*

International Joint Conference 15th International Conference on Computational Intelligence in Security for Information Systems (CISIS 2022) 13th International Conference on EUropean Transnational Education (ICEUTE 2022)

Proceedings

Springer

Lecture Notes in Networks and Systems

Volume 532

The series “Lecture Notes in Networks and Systems” publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

More information about this series at https://link.springer.com/bookseries/15179

Pablo García Bringas · Hilde Pérez García ·
Francisco Javier Martínez de Pisón ·
José Ramón Villar Flecha ·
Alicia Troncoso Lora · Enrique A. de la Cal ·
Álvaro Herrero · Francisco Martínez Álvarez ·
Giuseppe Psaila · Héctor Quintián ·
Emilio Corchado
Editors

International Joint Conference 15th International Conference on Computational Intelligence in Security for Information Systems (CISIS 2022) 13th International Conference on EUropean Transnational Education (ICEUTE 2022)

Proceedings

Springer

Editors
Pablo García Bringas
Faculty of Engineering
University of Deusto
Bilbao, Spain

Francisco Javier Martínez de Pisón
University of La Rioja
Logroño, Spain

Alicia Troncoso Lora
Data Science and Big Data Lab
Pablo de Olavide University
Sevilla, Spain

Álvaro Herrero
Departamento de Ingeniería Informática,
Escuela Politécnica Superior
University of Burgos
Burgos, Spain

Giuseppe Psaila
University of Bergamo
Bergamo, Italy

Emilio Corchado
University of Salamanca
Salamanca, Spain

Hilde Pérez García
University of León
León, Spain

José Ramón Villar Flecha
University of Oviedo
Oviedo, Spain

Enrique A. de la Cal
University of Oviedo
Oviedo, Spain

Francisco Martínez Álvarez
Pablo de Olavide University
Seville, Spain

Héctor Quintián
Department of Industrial Engineering
University of A Coruña
Ferrol, Spain

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-3-031-18408-6 ISBN 978-3-031-18409-3 (eBook)
https://doi.org/10.1007/978-3-031-18409-3

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume of Lecture Notes in Networks and Systems contains accepted papers presented at the 15th International Conference on Computational Intelligence in Security for Information Systems (CISIS 2022) and the 13th International Conference on EUropean Transnational Education (ICEUTE 2022). These conferences were held in the beautiful city of Salamanca, Spain, in September 2022.

The aim of CISIS 2022 is to offer a meeting opportunity for academic and industry-related researchers belonging to the various, vast communities of computational intelligence, information security, and data mining. The need for intelligent, flexible behavior by large, complex systems, especially in mission-critical domains, is intended to be the catalyst and the aggregation stimulus for the overall event.

After a through peer review process, the CISIS 2022 International Program Committee selected 20 papers, which are published in this conference proceedings. In this edition, three special sessions were organized: Cybersecurity in Future Connected Societies, Cybersecurity and Trusted Supply Chains of ICT and Intelligent Solutions for Cybersecurity Systems.

The aim of ICEUTE 2022 is to offer a meeting point for people working on transnational education within Europe. It provides a stimulating and fruitful forum for presenting and discussing the latest works and advances on transnational education within European countries.

In the case of ICEUTE 2022, the International Program Committee selected five papers, which are also published in this conference proceedings.

The selection of papers was extremely rigorous to maintain the high quality of the conferences. We want to thank the members of the program committees for their hard work during the reviewing process. This is a crucial process for creating a high-standard conference; CISIS and ICEUTE would not exist without their help.

CISIS 2022 and ICEUTE 2022 enjoyed outstanding keynote speeches by distinguished guest speakers: Prof. Ajith Abraham, Director of Machine Intelligence Research Labs (MIR Labs), and Prof. Guy De Tré, head of the research group on Database, Document, and Content Management (DDCM) at Ghent University, Belgium, and Félix Barrio Gen-eral Director at INCIBE (Spain).

CISIS 2022 has teamed up with "Logic Journal of the IGPL" (Oxford University Press), and Expert Systems (Wiley) for a suite of special issues, including selected papers from CISIS 2022.

Particular thanks go as well to the conference's main sponsors, Startup Olé, the CYL-HUB project financed with NEXT-GENERATION funds from the European Union, the Ministry of Labor and Social Economy, the Recovery, Transformation and Resilience Plan and the State Public Employment Service, channeled through the Junta de Castilla y León, BISITE research group at the University of Salamanca, CTC research group at the University of A Coruña, and the University of Salamanca. They jointly contributed in an active and constructive manner to the success of this initiative.

We would like to thank all the special session organizers, contributing authors, as well as the members of the program committees and the local organizing committee for their hard and highly valuable work. Their work has helped to contribute to the success of the CISIS and ICEUTE 2022 events.

September 2022

Pablo García Bringas
Hilde Pérez García
Francisco Javier Martínez de Pisón
José Ramón Villar Flecha
Alicia Troncoso Lora
Enrique A. de la Cal
Álvaro Herrero
Francisco Martínez Álvarez
Giuseppe Psaila
Héctor Quintián
Emilio Corchado

Organization

CISIS 2021

General Chair

Emilio Corchado	University of Salamanca, Spain

Program Committee Chairs

Pablo García Bringas	University of Deusto, Spain
Hilde Pérez García	University of León, Spain
Francisco Javier Martínez de Pisón	University of La Rioja, Spain
José Ramón Villar Flecha	University of Oviedo, Spain
Alicia Troncoso Lora	Pablo Olavide University, Spain
Enrique A. de la Cal	University of Oviedo, Spain
Álvaro Herrero	University of Burgos, Spain
Francisco Martínez Álvarez	Pablo Olavide University, Spain
Giuseppe Psaila	University of Bergamo, Italy
Héctor Quintián	University of A Coruña, Spain
Emilio Corchado	University of Salamanca, Spain

Program Committee

Adam Wójtowicz	Poznań University of Economics and Business, Poland
Agustin Martin Muñoz	C.S.I.C., Spain
Alberto Peinado	University of Malaga, Spain
Álvaro Herrero Cosio	University of Burgos, Spain
Álvaro Michelena Grandío	University of A Coruña, Spain
Amparo Fuster-Sabater	C.S.I.C., Spain

Anca Avram	Technical University of Cluj-Napoca, North University Center at Baia Mare, Romania
Andreea Vescan	Babes-Bolyai University, Cluj-Napoca, Romania
Andrysiak Tomasz	University of Technology and Life Sciences (UTP), Poland
Angel Arroyo	University of Burgos, Spain
Angel Martin Del Rey	University of Salamanca, Spain
Araceli Queiruga-Dios	University of Salamanca, Spain
Borja Sanz	University of Deusto, Spain
Carlos Cambra	University of Burgos, Spain
Carlos Pereira	ISEC, Portugal
Carmen Benavides	University of León, Spain
Ciprian Pungila	West University of Timișoara, Romania
Cristina Alcaraz	University of Malaga, Spain
Daniel Urda	University of Burgos, Spain
David Arroyo	C.S.I.C., Spain
David Garcia-Retuerta	University of Salamanca, Spain
Eduardo Solteiro Pires	UTAD University, Portugal
Elena Hernández Nieves	University of Salamanca, Spain
Emilia Cioroaica	Fraunhofer IESE, Germany
Eneko Osaba	TECNALIA Research & Innovation, Spain
Enrique Onieva	University of Deusto, Spain
Esteban Jove	University of A Coruña, Spain
Esteban Maurin	University of Salamanca, Spain
Fernando Ribeiro	EST, Portugal
Fernando Tricas	University of Zaragoza, Spain
Francisco Martínez-Álvarez	Pablo de Olavide University, Spain
Francisco Zayas-Gato	University of A Coruña, Spain
Guillermo Morales-Luna	CINVESTAV-IPN, Mexico
Héctor Quintián	University of A Coruña, Spain
Hugo Sanjurjo-González	University of Deusto, Spain
Hugo Scolnik	ARSAT SA, Argentina
Igor Santos	Mondragon Unibertsitatea, Spain
Ioannis Sorokos	Fraunhofer IESE, Germany
Isaias Garcia	University of León, Spain
Javier Nieves	Azterlan, Spain
Jesús Díaz-Verdejo	University of Granada, Spain
Jose A. Onieva	University of Malaga, Spain
Jose Barata	NOVA University Lisbon, Portugal
Jose Carlos Metrolho	IPCB, Portugal
José Francisco Torres Maldonado	Pablo de Olavide University, Spain
Jose Luis Calvo-Rolle	University of A Coruña, Spain
José Luis Casteleiro-Roca	University of A Coruña, Spain
José Luis Imaña	Complutense University of Madrid, Spain

Jose M. Molina	University Carlos III of Madrid, Spain
Jose Manuel Lopez-Guede	University of the Basque Country, Spain
Josep Ferrer	University of les Illes Balears, Spain
Juan J. Gude	University of Deusto, Spain
Juan Jesús Barbarán	University of Granada, Spain
Juan Pedro Hecht	University of Buenos Aires, Argentina
Lidia Sánchez-González	Universidad de León, Spain
Luis Alfonso Fernández Serantes	FH-Joanneum University of Applied Sciences, Spain
Luis Hernandez Encinas	C.S.I.C., Spain
Manuel Castejón-Limas	Universidad de León, Spain
Manuel Graña	University of the Basque Country, Spain
Marc Ohm	Universität Bonn, Germany
Marina Castaño Ishizaka	University of Salamanca, Spain
Michal Choras	ITTI, Poland
Michal Wozniak	Wroclaw University of Technology, Poland
Míriam Timiraos Díaz	University of A Coruña, Spain
Nuño Basurto	University of Burgos, Spain
Oliviu Matei	North University of Baia Mare, Romania
Oscar Llorente-Vazquez	University of Deusto, Spain
Ovidiu Cosma	Technical University Cluj-Napoca, Romania
Pablo García Bringas	University of Deusto, Spain
Petrica Pop	Technical University of Cluj-Napoca, North University Center at Baia Mare, Romania
Rafael Alvarez	University of Alicante, Spain
Rafael Corchuelo	University of Seville, Spain
Raúl Durán	University of Alcalá, Spain
Robert Burduk	University Wroclaw, Poland
Roberto Casado-Vara	University of Burgos, Spain
Rogério Dionísio	Polytechnic Institute of Castelo Branco, Portugal
Rudolf Erdei	Holisun SRL, Romania
Salvador Alcaraz	Miguel Hernandez University, Spain
Sorin Stratulat	Université de Lorraine, France
Wenjian Luo	Harbin Institute of Technology, China
Wilson Rojas	El Bosque University, Colombia

CISIS 2022: Special Sessions

Cybersecurity in Future Connected Societies

Program Committee

Koji Ishibashi (Organizer)	University of Electro-Communications, Japan
Víctor Gayoso Martínez (Organizer)	C.S.I.C., Spain

Kazuo Sakiyama	The University of Electro-Communications, Japan
Lorenzo Mucchi	Dept. of Information Engineering, University of Florence, Italy
Mitsugu Iwamoto	University of Electro-Communications, Japan
Naoki Ishikawa	Yokohama National University, Japan
Satyanarayana Vuppala	CITI Bank, Ireland
Seyedomid Motlagh	RWTH Aachen University, Germany
Slobodan Petrovic	Norwegian University of Science and Technology, Norway

Cybersecurity and Trusted Supply Chains of ICT

Program Committee

Anca Andreica (Organizer)	Babes-Bolyai University, Romania
Antonio Skarmeta Gomez (Organizer)	Universidad de Murcia, Spain
Camelia Chira (Organizer)	Babes-Bolyai University, Romania
Eda Marchetti (Organizer)	Isti-cnr, Italy
Ileana Buhan (Organizer)	RU, Netherlands
Jose Barata (Organizer)	UNINOVA, Portugal
Oliviu Matei (Organizer)	HOLISUN, Romania
Ovidiu Cosma (Organizer)	Technical University of Cluj-Napoca, Romania
Petrica Pop (Organizer)	Technical University of Cluj-Napoca, Romania
Ricardo Silva Peres (Organizer)	Universidade NOVA de Lisboa, Portugal

Intelligent Solutions for Cybersecurity

Program Committee

Bernabé Dorronsoro (Organizer)	University of Cádiz, Spain
Daniel Urda Muñoz (Organizer)	University of Burgos, Spain
Meelis Kull (Organizer)	University of Tartu, Estonia
Roberto Magán Carrión (Organizer)	University of Granada, Spain
Ángel Canal Alonso	University of Salamanca, Spain
Carlos Cambra	University of Burgos, Spain
Carlos Fernandes	Universidad de Lisboa, Portugal
Eduardo Feo Flushing	Carnegie Mellon University in Qatar, Qatar
Esther-LydiaSilva-Ramírez	University of Cádiz, Spain
Jamal Toutouh	Massachusetts Institute of Technology, USA

Javier González-Enrique	University of Cádiz, Spain
José Manuel García Nieto	University of Málaga, Spain
Lipika Deka	De Montfort University, UK
Margarita Robles Carrillo	University of Granada, Spain
Nuño Basurto	University of Burgos, Spain
Pablo García	University of Granada, Spain
Pablo Ruiz	OriGen AI, Spain
Rafael A. Rodríguez-Gómez	International University of La Rioja, Spain
Rafael M. Luque-Baena	University of Extremadura, Spain

CISIS 2022 Organizing Committee Chairs

Emilio Corchado	University of Salamanca, Spain
Héctor Quintián	University of A Coruña, Spain

CISIS 2022 Organizing Committee

Álvaro Herrero Cosio	University of Burgos, Spain
José Luis Calvo Rolle	University of A Coruña, Spain
Ángel Arroyo	University of Burgos, Spain
Daniel Urda	University of Burgos, Spain
Nuño Basurto	University of Burgos, Spain
Carlos Cambra	University of Burgos, Spain
Esteban Jove	University of A Coruña, Spain
José Luis Casteleiro Roca	University of A Coruña, Spain
Francisco Zayas Gato	University of A Coruña, Spain
Álvaro Michelena	University of A Coruña, Spain
Míriam Timiraos Díaz	University of A Coruña, Spain

ICEUTE 2022

Organization

General Chair

Emilio Corchado	University of Salamanca, Spain

Program Committee Chairs

Pablo García Bringas	University of Deusto, Spain
Hilde Pérez García	University of León, Spain
Francisco Javier Martínez de Pisón	University of La Rioja, Spain

José Ramón Villar Flecha	University of Oviedo, Spain
Alicia Troncoso Lora	Pablo Olavide University, Spain
Enrique A. de la Cal	University of Oviedo, Spain
Álvaro Herrero	University of Burgos, Spain
Francisco Martínez Álvarez	Pablo Olavide University, Spain
Giuseppe Psaila	University of Bergamo, Italy
Héctor Quintián	University of A Coruña, Spain
Emilio Corchado	University of Salamanca, Spain

Program Committee

Alessandra Raffaetà	Ca' Foscari University of Venezia, Italy
Álvaro Michelena Grandío	University of A Coruña, Spain
Ana Maria Lara Palma	University of Burgos, Spain
Ana Rosa Pereira Borges	ISEC-Coimbra Polytechnic Institute, Portugal
Andreea Vescan	Babes-Bolyai University, Romania
Angel Arroyo	University of Burgos, Spain
Antonio Morales-Esteban	University of Seville, Spain
Carlos Cambra	University of Burgos, Spain
Carlos Pereira	ISEC, Portugal
Carmen Benavides	University of León, Spain
Daniel Urda	University of Burgos, Spain
Daniela Zaharie	West University of Timisoara, Romania
David Becerra Alonso	University of Loyola, Spain
Diego Pedro Pinto Roa	Universidad Nacional de Asunción, Paraguay
Domingo S. Rodriguez Baena	Pablo de Olavide University, Spain
Dragan Simic	University of Novi Sad, Serbia
Eduardo Solteiro Pires	UTAD University, Portugal
Eloy Irigoyen	University of the Basque Country, Spain
Esteban Jove	University of A Coruña, Spain
Estibaliz Apiñaniz	University of the Basque Country, Spain
Federico Divina	Pablo de Olavide University, Spain
Francisco Gomez Vela	Pablo de Olavide University, Spain
Francisco Martínez-Álvarez	Pablo de Olavide University, Spain
Francisco Zayas-Gato	University of A Coruña, Spain
Guillermo Santamaria	CONACYT-INEEL, Mexico
Héctor Quintián	University of A Coruña, Spain
Hugo Sanjurjo-González	University of Deusto, Spain
Isaias Garcia	University of León, Spain
J. David Nuñez-Gonzalez	University of the Basque Country, Spain
Jairo Ortiz-Revilla	University of Burgos, Spain
Jorge Barbosa	ISEC, Portugal
José Francisco Torres Maldonado	Pablo de Olavide University, Spain

Jose Luis Calvo-Rolle	University of A Coruña, Spain
José Luis Casteleiro-Roca	University of A Coruña, Spain
Jose Luis Vazquez Noguera	Universidad Nacional de Asunción, Paraguay
José Manuel Galán	University of Burgos, Spain
Jose Manuel Lopez-Guede	University of the Basque Country, Spain
José-Lázaro Amaro-Mellado	University of Seville, Spain
Juan J. Gude	University of Deusto, Spain
Juan Pavón	Complutense University of Madrid, Spain
Julián Estévez	University of the Basque Country, Spain
Laura Fernández-Robles	University of León, Spain
Laura Melgar-García	Pablo de Olavide University, Spain
Lidia Sánchez-González	University of León, Spain
Luis Alfonso Fernández Serantes	FH-Joanneum University of Applied Sciences, Spain
Manuel Castejón-Limas	University of León, Spain
María Fernanda Gambarini	Francisco de Vitoria University, Spain
Maria Jose Marcelino	University of Coimbra, Portugal
Maria Teresa Godinho	Polytechnic Institute of Beja, Portugal
Maria Victoria Requena	University of Seville, Spain
Marián Queiruga-Dios	Francisco de Vitoria University, Spain
Miguel Ángel Queiruga-Dios	University of Burgos, Spain
Miguel Carriegos	RIASC, Spain
Miguel Garcia Torres	Pablo de Olavide University, Spain
Míriam Timiraos Díaz	University of A Coruña, Spain
Nuño Basurto	University of Burgos, Spain
Pablo García Bringas	University of Deusto, Spain
Paola Clara Leotta	University of Catania, Italy
Paulo Moura Oliveira	UTAD University, Portugal
Pedro Mauricio Acosta Castellanos	Santo Tomás University, Colombia
Ramona Mihaila	Dimitrie Cantemir Christian University, Romania
Richard Duro	University of A Coruña, Spain
Sebastian Alberto Grillo	Universidad Autonoma de Asunción, Paraguay
Soraya Muñoz Pérez	University Francisco de Vitoria, Spain
Sorin Stratulat	Université de Lorraine, France
Viviana Elizabeth Jiménez Chaves	Universidad Nacional de Itapúa, Paraguay

ICEUTE 2022 Organizing Committee Chairs

Emilio Corchado	University of Salamanca, Spain
Héctor Quintián	University of A Coruña, Spain

ICEUTE 2022 Organizing Committee

Álvaro Herrero Cosio	University of Burgos, Spain
José Luis Calvo Rolle	University of A Coruña, Spain
Ángel Arroyo	University of Burgos, Spain
Daniel Urda	University of Burgos, Spain
Nuño Basurto	University of Burgos, Spain
Carlos Cambra	University of Burgos, Spain
Esteban Jove	University of A Coruña, Spain
José Luis Casteleiro Roca	University of A Coruña, Spain
Francisco Zayas Gato	University of A Coruña, Spain
Álvaro Michelena	University of A Coruña, Spain
Míriam Timiraos Díaz	University of A Coruña, Spain

Contents

CISIS Applications

Analysis of Long-Range Forecast Strategies for IoT on Urban Water Consumption Prediction Task 3
Krzysztof Pałczyński, Tomasz Andrysiak, Marcin Głowacki, Michał Kierul, and Tomasz Kierul

Genetic Algorithm Based Aggregation for Federated Learning in Industrial Cyber Physical Systems 12
Souhila Badra Guendouzi, Samir Ouchani, and Mimoun Malki

Hand SOS Gesture Detection by Computer Vision 22
Roberto Viejo-López, Virginia Riego del Castillo, and Lidia Sánchez-González

Prediction of Smart Energy Meter Network Traffic Features for Anomaly Detection 30
Łukasz Saganowski and Tomasz Andrysiak

An Anomaly Detection Approach for Realtime Identification Systems Based on Centroids 40
Álvaro Michelena, Francisco Zayas-Gato, Esteban Jove, José-Luis Casteleiro-Roca, Héctor Quintián, Óscar Fontenla-Romero, and José Luis Calvo-Rolle

Powerful Biogeography-Based Optimization Algorithm with Local Search Mechanism for Job Shop Scheduling Problem with Additional Constraints 52
Madiha Harrabi, Olfa Belkahla Driss, and Khaled Ghedira

Dimensionality-Reduction Methods for the Analysis of Web Traffic . . . 62
Nuño Basurto, Álvaro Michelena, Daniel Urda, Hector Quintián, José Luis Calvo-Rolle, and Álvaro Herrero

Special Session on Cybersecurity in Future Connected Societies

About the Fujisaki-Okamoto Transformation in the Code-Based Algorithms of the NIST Post-quantum Call 75
Miguel Ángel González de la Torre and Luis Hernández Encinas

Analysis of Secret Key Agreement Protocol for Massive MIMO Systems 86
Seiya Otsuka, Hiroki Iimori, Kengo Ando, Giuseppe Thadeu Freitas de Abreu, Koji Ishibashi, and Naoki Ishikawa

Efficient Implementation of Stream Cipher SNOW 3G for Resource-Constrained Devices 95
G. Cotrina, A. Peinado, and A. Ortiz

State of the Art of Cybersecurity in Cooperative, Connected and Automated Mobility 104
Óscar Castillo Campo, Víctor Gayoso Martínez, Luis Hernández Encinas, Agustín Martín Muñoz, and Roberto Álvarez Fernández

Cryptographic Protocols in Advanced Metering Infrastructures in Smart Grids 114
Luis Hernández-Álvarez, Juan José Bullón Pérez, and Araceli Queiruga-Dios

Special Session on Cybersecurity and Trusted Supply Chains of ICT

Orchestrator Architecture and Communication Methodology for Flexible Event Driven Message Based Communication 127
Rudolf Erdei, Daniela Delinschi, Emil Pașca, and Oliviu Matei

A Comparative Study of Machine Learning Algorithms for the Detection of Vulnerable Python Libraries 138
Laura Pérez-Vilarelle, Eva Sotos Martínez, and Javier Yépez Martínez

Evaluation of the Reliability Index of IP Addresses in Reputation Lists 149
Alberto Miranda-García, Ignacio Samuel Crespo-Martínez, Ángel Manuel Guerrero-Higueras, and Vicente Matellán-Olivera

Forecasting the Number of Bugs and Vulnerabilities in Software Components Using Neural Network Models 159
Ovidiu Cosma, Petrică Pop, Cosmin Sabo, and Laura Cosma

Special Session on Intelligent Solutions for Cybersecurity Systems

Reinforcement Learning Model Free with GLIE Monte-Carlo on Policy Update for Network Topology Discovery 171
Roberto Casado-Vara, Marcos Severt, Ángel Martín del Rey, Héctor Quintián, and Jose L. Calvo-Rolle

Obfuscating LLVM Intermediate Representation Source Code with NSGA-II 181
Juan Carlos de la Torre, José Miguel Aragón-Jurado, Javier Jareño, Sébastien Varrette, and Bernabé Dorronsoro

A Deep Learning-Based Approach for Mimicking Network Topologies: The Neris Botnet as a Case of Study 192
Francisco Álvarez-Terribas, Roberto Magán-Carrión, Gabriel Maciá-Fernández, and Antonio M. Mora García

ICEUTE

Evaluating Classifiers' Performance to Detect Attacks in Website Traffic 205
Daniel Urda, Nuño Basurto, Meelis Kull, and Álvaro Herrero

Evaluation of an Interactive Guide for Robotics Self-learning 216
Álvaro Ovejero, Jesús Enrique Sierra-García, and Matilde Santos

Gamifying the Classroom for the Acquisition of Skills Associated with Machine Learning: A Two-Year Case Study 224
Antonio M. Durán-Rosal, David Guijo-Rubio, Víctor M. Vargas, Antonio M. Gómez-Orellana, Pedro A. Gutiérrez, and Juan C. Fernández

Hackathon in Teaching: Applying Machine Learning to Life Sciences Tasks 236
David Guijo-Rubio, Víctor M. Vargas, Javier Barbero-Gómez, Jose V. Die, and Pablo González-Moreno

Digital Platforms for Education. The Case of e4you 247
Javier Parra-Domínguez, Sergio Manzano, Susana Herrero, and Pablo Chamoso

3D Virtual Laboratory for Control Engineering Using Blended Learning Methodology 256
Francisco Zayas-Gato, Álvaro Michelena, Esteban Jove, José-Luis Casteleiro-Roca, Héctor Quintián, Elena Arce, and José Luis Calvo-Rolle

Author Index 267

CISIS Applications

Analysis of Long-Range Forecast Strategies for IoT on Urban Water Consumption Prediction Task

Krzysztof Pałczyński[1](✉), Tomasz Andrysiak[1], Marcin Głowacki[1], Michał Kierul[2], and Tomasz Kierul[2]

[1] Bydgoszcz University of Science and Technology, Bydgoszcz, Poland
{krzpal004,andrys,marglo010}@pbs.edu.pl
[2] Research and Development Center SOFTBLUE S.A., Bydgoszcz, Poland
{mkierul,tkierul}@softblue.pl

Abstract. With the rapid development of technology, researchers worldwide have applied the Internet of Things to effectively transmit and monitor water levels and detect anomalies in real time. The data obtained enables numerical methods to predict water consumption as well. In the presented paper, an attempt has been made to predict water consumption for various problems forward using dedicated models and a system using an iterative approach. For this purpose, neural network algorithms such as Random Forest, XGBoost, Decision Tree, and Support Vector Regression were tested and used to train the prediction models. The results presented allowed to indicate the difference between the examined methods through the Mean Absolute Percentage Error of prediction. The used set of algorithms allowed to show the problem of estimating water prediction from different points of view. Thus, determining the tested systems' seasonality and short-term and long-term trends. This allowed to choose the two best algorithms, one that needs less computational power to work; this seems to be a better solution.

Keywords: Prediction systems · Machine learning · Water consumption forecast

1 Introduction

The issues involved in managing water distribution networks are based on a large number of processes. The demand for water for service or residential purposes affects the amount of water withdrawn. In both cases, it is necessary to estimate consumption based on the services provided by the company or the inhabitants' behavior. The topic of demand analysis is essential due to the number of variables depending on the type of facility to which water is supplied, which provides the basis for the proper design of water management strategies. Forecasts are of great importance in the context of decisions related to introducing procedures to optimize water treatment plants or pumping stations.

P. García Bringas et al. (Eds.): CISIS 2022/ICEUTE 2022, LNNS 532, pp. 3–11, 2023.
https://doi.org/10.1007/978-3-031-18409-3_1

In the presented work, an attempt was made to predict water consumption forward by different numbers of probes. The water consumption signal under study was divided into 15-min intervals, with one step covering the 15 min elapsed. For the long-range prediction calculations, systems were used that could attempt to estimate water consumption using two methods; using specially trained models for forecasting each step into the future and using only one model trained for next-value prediction and using it iteratively to find demanded forecasts. The first involved a dedicated model designed to look one, two, or several time steps ahead. The second worked based on performing predictions using earlier forecasting results. This system used an iterative approach involving only one model, with each forward time step depending on the previous one. It goes forward when the prediction turns out to be true after predicting the following sample. The comparison between systems was established based on their mean errors and suitability for Internet of Things (IoT) purposes.

The thesis in this article is that long-range prediction obtained from the chain of short-range predictions made using the results of previous forecasts is an acceptable predictive strategy for specific time ranges. IoT applications would benefit from having only one trained model and using it for long-range predictions instead of keeping many models, each trained for forecasting on its specific time range.

This paper aims to use selected algorithms to train forecasting models to predict the following sample and evaluate their performance using Mean Absolute Percentage Error (MAPE). Short and long-term forecast results of all tested algorithms are presented, highlighting the performance of iterative systems in short-term and dedicated models in long-term forecasts.

The data came from the waterworks and was taken thanks to a telemetry overlay on the water meters using IoT. In simple terms, it is a system of electronic devices that can automatically communicate and exchange data over a network without human intervention. IoT-based solutions in water level monitoring and storage are a novelty, as well as the analysis of long-range forecasting strategies suited for low-computational power devices commonly used for IoT purposes. The proper use of the received data can help to reduce the occurrence of anomalies and mitigate the possible effects of a water crisis [1].

This paper is organized as follows: an introduction, Sect. 2, in which work on water use forecasting to detect anomalies and determine the causes of fluctuations in consumption is discussed. The third section includes a detailed presentation of the systems used and the rationale for choosing particular algorithms. Section 4 discusses the experimental results. The paper concludes with a sub-summary.

2 Related Work

Several approaches to the problem of water demand forecasting, or drought forecasting, can be found in the literature. Typically, statistical models have been used to estimate losses and detect undesirable anomalies.

In paper [2], an algorithm was tested to detect and locate water leaks at an early stage of demand forecasting. The algorithm used was based on pattern similarity, and the demonstration area of the project was a medium-sized city in Spain. The algorithm was tested in three different scenarios over two years (2014–2016). The scenarios were selected based on other profiles, including industrial areas, city centers, and suburbs. The prediction analysis yielded good results. The predictions indicated the most accurate results for the downtown area due to the many people living there. The authors suggest that the method is suitable for predicting daily water flow. The signals proved highly unpredictable for the time range studied due to water flow spikes caused by anomalous days such as holidays. As the historical database of similar past situations increases, it is possible to increase the prediction accuracy. The task and problems related to water flow spikes are identical to the subject of our work. However, our analyses focused on the approaches to long-term prediction and divergence rate of forecast from the base signal.

Another way of determining the unit water demand and consumption distribution was addressed in work [3], which focused on the hourly analysis of water consumption in selected residential buildings after declaring an epidemic emergency in the country. The objective was to verify the application of clustering using mean class clustering. The research objects were three similar five-story residential buildings in the same housing estate in Bydgoszcz. The study time included 464 measurement days, of which 276 days covered the period before the COVID-19 pandemic and 188 during the pandemic, which was included in the calculations. The application of cluster analysis through clustering by the mean squared method allowed us to develop a pattern of hourly water demand by working and non-working days. The results clearly show the profound changes in daily water demand patterns exerted by the COVID-19 pandemic. Continuous monitoring is needed to obtain up-to-date consumption patterns. The averaged histograms obtained can be used to determine nodal water consumption in the mathematical modeling of water supply networks. The data analyzed in this work is also from times of the COVID pandemic. However, in this work, the predictions were conducted using machine learning algorithms and Deep Learning.

Another interesting paper [4] addressed clustering and Support Vector Regression for water demand forecasting and the possibility of anomaly detection. The data used for the analysis was obtained from a municipal water company in Milan, Italy, from a Supervisory Control And Data Acquisition (SCADA) system and 26 sensors, i.e., meters used at the individual level. The proposed approach offers several innovations, from increasing the forecasting accuracy at the municipal and individual classes to using a time series clustering procedure on the raw data. The second innovation is to learn different Support Vector Machine (SVM) regression models separately for each cluster and each hour of water consumption. This results in the optimization of the pumping schedule and the ability to identify online anomalies resulting from wise meter faults, cyber-physical attacks, or changes in habits. The proposed approach has obtained favorable results both in the municipal application and by individual consumers. The SVR model was also employed in our work. However, the preliminary results for our data set did not indicate any benefit in dividing water consumption signals into specific periods.

In [5], the authors attempted to change the operating conditions of water supply systems using empirical exponents. The focus was mainly on modifying the flow rate and hydrostatic pressure. Many failures were observed in managing the water supply network's stress. The study area covered the Municipal Water Supply and Sewerage Company S.A. in Wrocław, and the recording time was between April 2017 and January 2018. The conducted research enabled the determination of an exponent linking the flow rate-pressure relationship to the overall assessment of the water supply network. Changes in the form of a steady, cascading decrease in pressure through the night flow period and a cascading increase in the morning resulted in the implementation of new analytical procedures for pressure and flow rate analysis. This work has a similar subject to this one. However, it is based on theoretical models and expert knowledge. This paper tackles the problem from a data science point of view and aims to define approaches that may be useful outside this particular problem involving water consumption forecasting.

In work [6], an analysis of water consumption variability based on combined long-term and short-term data was performed. The aim of the work was based on the development of a new deterministic method that would allow the data from periodic measurements to be collected together. To then serially determine the degree of variation in water consumption. The object of study was three blocks of flats located in Świdnica, and the observation time was five years. The experiment results show that it is impossible to analyze the cumulative fluctuations of water consumption for data without preprocessing. To categorize the measurements, it is necessary to supplement the data sets with water consumption records for several sufficiently long periods. Research conducted in our paper also deemed pre-processing necessary as water consumption signal tends to have random spikes of activity, which disrupts mean squared error-based loss function driven neural networks. To perform the experiments, we had to draw the same conclusions.

On the other hand, in the paper [7], seasonal forecasts were analyzed, and the main interest was developing a drought forecast. Drought in 1998–2002, which occurred in South Carolina areas, is considered the most severe problem for the state. The analysis performed was extended to include other drought indicators, and the information to calculate these indicators is provided every month in terms of precipitation and temperature. Thus, different scenarios can be adapted to determine the amount of rain needed to end drought conditions and the probability of receiving those amounts. The article bases its methodology on bootstrapping, which we also used in our Decision Tree-based algorithms and can confirm its efficiency in forecasting.

3 Methodology

3.1 Dataset

The dataset used in this research was obtained from the measurements of the pump room, generating an average water supply of 8 m^3 per hour. The data is the property of Softblue S.A company. Access to the data was provided to conduct this research. The signal was initially encoded as a timestamp between subsequent transfers of 10 L of water through the pump house. For this research, the signal was resampled to contain the amount of water transferred every 15 min. Due to temporal anomalies in water supply provision, data cleaning was required. Each value of material water supply exceeded the 94 percentile of water flow distribution and was reduced to the value associated with the 94 percentile.

3.2 Experiment Design

The cleaned data are formed into observation vectors. Each observation vector contains the last 20 values representing the previous 5 h of the observed water consumption. Two forecasting systems use these data. The first system uses models explicitly trained to predict a fixed amount of future samples. For example, the system that can predict the next 1, 2, 3 … 6 samples in the future contains six models, each trained to forecast its respective amounts of samples in the future. The second system has only one model trained to predict the following sample. The system uses this model to predict an arbitrarily chosen number of samples in the future by iteratively predicting the following sample and uses the value of the prediction to update the observation vector for another sample prediction. The systems' algorithms are presented in Fig. 1.

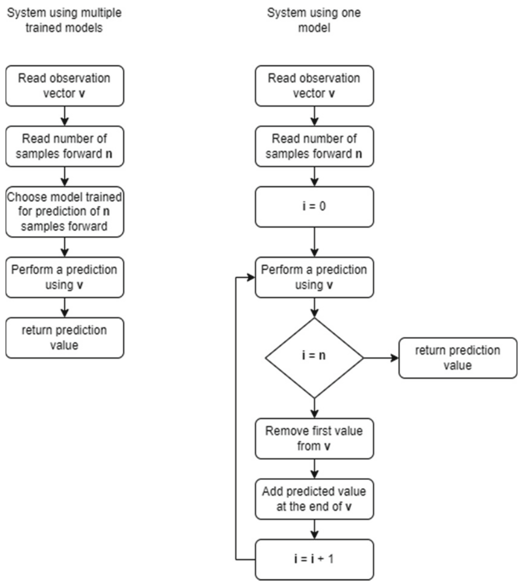

Fig. 1. Presentation of the algorithms forming prediction systems evaluated in this work. On the left, the system using dedicated models is presented. On the right, the system employing an iterative approach is described.

3.3 Machine Learning Algorithms

Each system employs machine learning models to forecast the following sample. The algorithms selected for training prediction models in this study are: neural networks with two fully connected layers (referend further in the text as FC2), recurrent neural networks [8] with two recurrent layers and one fully connected (referred further in the text as RNN), Random Forest [9], XGBoost [10], Decision Tree [11] and Support Vector Regression [12] (SVR).

These algorithms were chosen due to their different approach to the forecasting problem. FC2 and SVR models perform the linear transformation of the input vector to obtain the forecast. RNN attempts to extract temporal patterns. The Decision Tree is the algorithm that divides solution space into subsets based on the partition purity. Random Forest and XGBoost perform an ensemble of decision trees in different ways. Such a collection of algorithms presents the forecasting problem from different angles and allows comparison of varying decision-making philosophies.

The critical factor to be considered during model selection is its ability to be implemented into the hardware commonly used in IoT applications. The machine learning algorithms (Random Forest, XGBoost, Decision Tree, and SVR) are implemented in the scikit-learn library widely used in Data Science. This library is lightweight enough to be used on the higher end of IoT devices like Raspberry microcomputers. Although frameworks for Deep Neural Networks may be too computationally expensive for IoT devices, the Fully Connected Neural Networks are algorithms simple enough to be implemented with reduced NumPy computational stack for micro-Python booted on the microcontrollers like ESP.

4 Experimental Results

The algorithms and techniques were evaluated using the Mean Absolute Percentage Error (MAPE) metric given by Eq. 1. The purpose of this metric is to present relative differences between the observed values predicted by the systems.

$$MAPE = \sum_{i=0}^{|Y|} \frac{\left|\widehat{Y_i} - Y_i\right|}{Y_i} \tag{1}$$

where Y denotes the vector of observed values and is the vector of predicted values.

The models were evaluated by averaging MAPE results from the 30 iterations of the experiment conducted for every triplet of the model algorithm, version of the system, and several samples to look into in the future. During each iteration, train, validation, and test datasets were randomly chosen from the available dataset.

The results are presented in Tables 1 and 2. The values were obtained from evaluations conducted on test datasets. In each table, rows contain values for tested models, and columns present models' performance on the number of samples of the future forecast. Table 1 contains values from the models dedicated to predicting a specific number of samples forward. Table 2 presents the evaluation results of the system using an iterative approach.

Table 1. MAPE metric values of the system using models dedicated to forecasting the fixed number of samples ahead. The columns denote the number of samples the model looks ahead.

Model	Sample count between last observed value and forecasted					
	1	2	3	4	5	6
SVR	9.4%	11.1%	12.1%	13.2%	14.6%	15.9%
RNN	9.4%	11.2%	12.2%	13.3%	14.4%	15.4%
Random forest	9.5%	11.0%	11.8%	12.8%	13.8%	15.0%
FC 2	9.9%	12.1%	13.7%	15.4%	17.1%	19.5%
XGBoost	10.0%	11.6%	12.4%	13.4%	14.5%	15.5%
Decision tree	13.7%	15.9%	16.9%	18.1%	19.3%	20.7%

Table 2. MAPE metric values of the system using one trained model iteratively. The columns denote the number of samples the model looks ahead.

Model	Sample count between last observed value and forecasted					
	1	2	3	4	5	6
SVR	9.4%	10.4%	13.0%	16.3%	19.9%	23.7%
RNN	9.4%	11.0%	13.7%	17.3%	21.6%	25.8%
Random forest	9.5%	10.6%	13.2%	16.1%	19.9%	23.6%
FC 2	9.9%	11.5%	14.3%	17.9%	21.8%	26.0%
XGBoost	10.0%	11.2%	13.7%	16.7%	20.2%	24.2%
Decision tree	13.7%	15.8%	18.4%	21.5%	24.8%	29.1%

According to the data stored in Tables 1 and 2, two algorithms with the best MAPE score on both the short and long-range forecasts are SVR and Random Forest. Figure 2 depicts the differences between the system using the dedicated model for forecasting and its iterative counterpart for both algorithms.

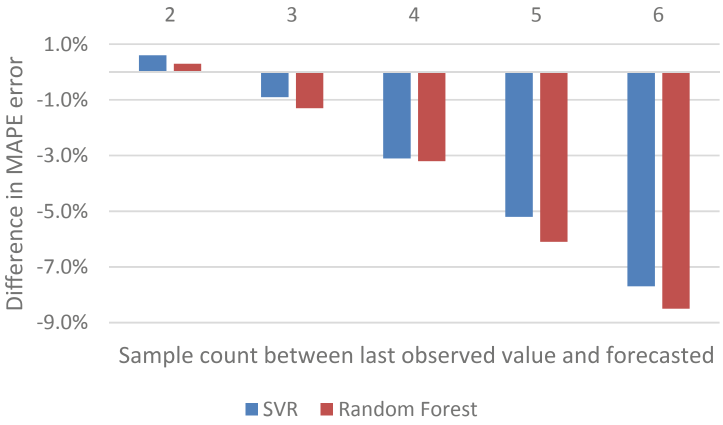

Fig. 2. Difference between MAPE error scores between systems using dedicated models for forecasting and their iterative counterparts.

5 Conclusion

Surprisingly, the iterative system proved slightly better than the system using dedicated models across all modeling algorithms. The improvement of the score ranges from 0.1% to 0.7%. Although the difference may not be considered significant, the repeating tendency on all tested algorithms may indicate that the water consumption signal measured in 15 min intervals may have a short-range trend in addition to its evident seasonality. In such a case, an iterative approach may stabilize the prediction, improving its average performance, albeit not for much.

Iterative systems' performance quickly deteriorates from fourth sample forward forecasting compared to the model with dedicated models. Assessment of further prediction is pointless due to error difference reaching almost nine percentage points. However, the iterative approach may be helpful for a relatively reliable long-range forecast to the fourth sample forward, for example, as a basis for anomaly detection systems. For such cases, the reduced number of models is preferable to better precision due to the limited memory of low-computational devices commonly used for IoT purposes. The thesis of this work has been proven suitable for forecasting up to four samples forward.

Systems with dedicated models provide a stable increase in the error of about one percentage point per number of forecast samples. The two best algorithms for this particular task are Random Forest and SVR. Their performance was the best for both short and long-range forecasts. However, SVR is an algorithm of much less computational complexity and memory dependence than Random Forest, resulting in a better solution for low-power devices like microcomputers.

Despite its more straightforward structure, the SVR model proved to have significantly better results than Neural Network with two fully connected layers (FC2). Overfitting is unlikely because a Recurrent Neural Network with two RNN layers obtained better results than FC2 despite a more complex structure. The margin function employed in SVR may be responsible for the improvement of the model. The margin function reduces the impact of the outliers by associating the constant penalty for values outside