

# Special Issue on Advances in Database Engineered Applications

Richard Chbeir <sup>1</sup>, Mirjana Ivanovic <sup>2</sup>, Yannis Manolopoulos <sup>3,\*</sup> and Claudio Silvestri <sup>4</sup>

<sup>1</sup> Department of Computer Science, University of Pau and the Adour Region, 64600 Anglet, France; richard.chbeir@univ-pau.fr

<sup>2</sup> Faculty of Sciences, University of Novi Sad, 21000 Novi Sad, Serbia; mira@dmi.uns.ac.rs

<sup>3</sup> Faculty of Pure and Applied Sciences, Open University of Cyprus, 2220 Nicosia, Cyprus

<sup>4</sup> Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University of Venice, 30172 Mestre, Italy; silvestri@dais.unive.it

\* Correspondence: yannis.manolopoulos@ouc.ac.cy

The 27th International Database Engineering and Applications Symposium (IDEAS-2023) was held in Heraklion, Crete, Greece, on 5–7 May 2023. This conference is an international forum for data engineering researchers, practitioners, developers, and application users to explore revolutionary ideas and results, and to exchange techniques, tools, and experiences. Following the long history of IDEAS events, proceedings have been published by ACM Press, plus Special Issues are being published with selected-invited papers. More information can be found on the conference website: <https://conferences.sigappfr.org/ideas2023/>, accessed on 11 March 2024.

The authors of a number of high-quality full papers were invited after the conference to submit revised and extended versions of their originally accepted conference papers to this Special Issue of *Computers*, published by MDPI, in an open access format. Each submission to this Special Issue should contain at least 50% new material, e.g., in the form of technical extensions, more in-depth evaluations, or additional use cases, and a change in the title, abstract, and keywords. These extended submissions underwent a peer-review process according to the journal's rules of action. At least two technical committees acted as reviewers for each extended article submitted to this Special Issue; in some cases, additional external reviewers were invited to guarantee a high-quality reviewing process. The present Special Issue contains six papers, which are briefly introduced as follows.

Contribution 1 by Simon Lohmann and Dietmar Tutsch (University of Wuppertal) examines the doubly linked tree of singly linked rings to provide hard real-time database operations on an FPGA. The authors present a hardware data structure specifically designed for FPGAs that enables the execution of the hard real-time database CRUD operations using a hybrid data structure that combines trees and rings. While the number of rows and columns has to be limited for hard real-time execution, the actual content can be of any size. Their structure restricts full navigational freedom to everything but the leaf layer, thus keeping the memory overhead for the data stored in the leaves low. Although its nodes differ in function, they all have exactly the same size and structure, reducing the number of cascaded decisions required in database operations. This enables fast and efficient hardware implementation on FPGAs. In addition to the usual comparison with known data structures, they also analyze the tradeoff between the memory consumption of our approach and a simplified version that is doubly linked in all layers.

Contribution 2 by Rohit Mittal, Geeta Rani (Manipal University Jaipur), Vibhakar Pathak (Arya College of Engineering and Information Technology, Jaipur), Sonam Chhikara, Vijaypal Singh Dhaka (Manipal University Jaipur), Eugenio Vocaturo, and Ester Zumpano (University of Calabria) examines low-cost multisensory robots for optimized path planning in diverse environments. The automation industry faces the challenge of avoiding interference with obstacles, estimating the next move of a robot, and optimizing its path in various environments. Although researchers have predicted the next move of a robot in



**Citation:** Chbeir, R.; Ivanovic, M.; Manolopoulos, Y.; Silvestri, C. Special Issue on Advances in Database Engineered Applications. *Computers* **2024**, *13*, 77. <https://doi.org/10.3390/computers13030077>

Received: 11 March 2024

Accepted: 12 March 2024

Published: 14 March 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

linear and non-linear environments, there is a lack of precise estimation of sectorial error probability while moving a robot on a curvy path. Additionally, existing approaches use visual sensors, incur high costs for robot design, and are ineffective in achieving motion stability on various surfaces. To address these issues, the authors propose a low-cost and multisensory robot capable of moving on an optimized path in diverse environments with eight degrees of freedom. They use the extended Kalman filter and an unscented Kalman filter for localization and position estimation of the robot. They also compare the sectorial path prediction error at different angles from  $0^\circ$  to  $180^\circ$  and demonstrate the mathematical modeling of various operations involved in navigating the robot. The minimum deviation of 1.125 cm between the actual and predicted paths proves the effectiveness of the robot in a real-life environment.

Contribution 3 by Vaclav Skala (University of West Bohemia, Pilsen) and Eliska Mourycova (University of Defence, Brno) focuses on meshfree interpolation of multi-dimensional time-varying scattered data. Interpolating and approximating scattered scalar and vector data is fundamental to resolving numerous engineering challenges. These methodologies predominantly rely on establishing a triangulated structure within the data domain, typically constrained to 2D or 3D dimensions. Subsequently, an interpolation or approximation technique is employed to yield a smooth and coherent outcome. The above article introduces a meshless methodology founded upon radial basis functions. This approach exhibits a nearly dimensionless character, facilitating the interpolation of data evolving over time. Specifically, it enables the interpolation of dispersed spatio-temporally varying data, allowing for interpolation within the space-time domain devoid of the conventional “time-frames”. Meshless methodologies tailored for scattered spatio-temporal data hold applicability across a spectrum of domains, encompassing the interpolation, approximation, and assessment of data originating from various sources, such as buoys, sensor networks, tsunami monitoring instruments, chemical and radiation detectors, vessel and submarine detection systems, weather forecasting models, as well as the compression and visualization of 3D vector fields, among others.

Contribution 4 by Benjamin Warnke, Stefan Fischer, and Sven Groppe (University of Luebeck) uses machine learning and routing protocols to optimize distributed SPARQL queries. Due to increasing digitization, the amount of data on the Internet of Things (IoT) is constantly increasing. To process queries efficiently, strategies must be found to reduce the transmitted data as much as possible. SPARQL is particularly well-suited to the IoT environment because it can handle various data structures. However, due to the flexibility of data structures, more data have to be joined again during processing. Therefore, a good join order is crucial, as it significantly impacts the number of intermediate results. Also, computing the best linking order is NP-hard because the total number of possible linking orders increases exponentially with the number of inputs, plus there are different definitions of optimal join orders. Machine learning uses stochastic methods to achieve good results even with complex problems quickly. Other DBMSs also consider reducing network traffic but neglect the network topology. Network topology is crucial in IoT as devices are not evenly distributed. In this respect, the authors present new techniques for collaboration between routing, application, and machine learning. Their approach, which pushes the operators as close as possible to the data source, minimizes the generated network traffic by 10%. Additionally, the model can reduce the number of intermediate results by a factor of 100 in comparison to other state-of-the-art approaches.

Contribution 5 by Giacomo Bergami, Samuel Appleby, and Graham Morgan (Newcastle University) examines specification mining over temporal data. Current specification mining algorithms for temporal data rely on exhaustive search approaches, which become detrimental in real data settings where a plethora of distinct temporal behaviors are recorded over prolonged observations. The above article proposes a novel algorithm, Bolt2, based on a refined heuristic search of our previous algorithm, Bolt. Their experiments show that the proposed approach not only surpasses exhaustive search methods in terms of running time but also guarantees a minimal description that captures the overall temporal

behavior. This is achieved through a hypothesis lattice search that exploits support metrics. Their novel specification mining algorithm also outperforms the results achieved in our previous contribution.

Contribution 6 by László Göcs and Zsolt Csaba Johanyák (John von Neumann University, Kecskemét) focuses on the feature selection problem with weighted ensemble ranking to improve the classification performance. Feature selection is a crucial step in machine learning, aiming to identify the most relevant features in high-dimensional data to reduce the computational complexity of model development and improve generalization performance. Ensemble feature-ranking methods combine the results of several feature-selection techniques to identify a subset of the most relevant features for a given task. In many cases, they produce a more comprehensive ranking of features than the individual methods used alone. The above article presents a novel approach to ensemble feature ranking, which uses a weighted average of the individual ranking scores calculated using these individual methods. The optimal weights are determined using a Taguchi-type design of experiments. The proposed methodology significantly improves classification performance on the CSE-CIC-IDS2018 dataset, particularly for attack types where traditional average-based feature-ranking score combinations result in low classification metrics.

The editors of this Special Issue would like to thank the authors and reviewers, as well as the editorial office of the journal *Computers*. This Special Issue would not have become a reality without their contributions and assistance.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflicts of interest.

#### List of Contributions:

1. Lohmann, S.; Tutsch, D. The Doubly Linked Tree of Singly Linked Rings: Providing Hard Real-Time Database Operations on an FPGA. *Computers* **2024**, *13*, 8. <https://doi.org/10.3390/computers13010008>.
2. Mittal, R.; Rani, G.; Pathak, V.; Chhikara, S.; Singh Dhaka, V.; Vocaturo, E.; Zumpano, E. Low-Cost Multisensory Robot for Optimized Path Planning in Diverse Environments. *Computers* **2023**, *12*, 250. <https://doi.org/10.3390/computers12120250>.
3. Skala, V.; Mourycova, E. Meshfree Interpolation of Multidimensional Time-Varying Scattered Data. *Computers* **2023**, *12*, 243. <https://doi.org/10.3390/computers12120243>.
4. Warnke, B.; Fischer, S.; Groppe, S. Using Machine Learning and Routing Protocols for Optimizing Distributed SPARQL Queries in Collaboration. *Computers* **2023**, *12*, 210. <https://doi.org/10.3390/computers12100210>.
5. Bergami, G.; Appleby, S.; Morgan, G. Specification Mining over Temporal Data. *Computers* **2023**, *12*, 185. <https://doi.org/10.3390/computers12090185>.
6. Göcs, L.; Csaba Johanyák, Z. Feature Selection with Weighted Ensemble Ranking for Improved Classification Performance on the CSE-CIC-IDS2018 Dataset. *Computers* **2023**, *12*, 147. <https://doi.org/10.3390/computers12080147>.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.