



Advances in Databases and Information Systems

Ladjet Bellatreche¹ · Patrick Valduriez² · Tadeusz Morzy³

© Springer Science+Business Media, LLC, part of Springer Nature 2017

1 Introduction

The success stories of the database technology keep companies continuously demanding more and more efficient services related to two principal entities: *data* and *queries*. The services dedicated to data concern mainly the following tasks: collecting, cleaning, filtering, integration, sharing, storing, transferring, visualizing, analyzing, securing. Regarding queries, services are often concentrated on processing, optimizing, personalizing and recommending. These services of both entities have to be revisited and extended to deal with the dimensions brought by Big Data Era and the new requirements of companies in the contexts of globalization, competition and climate change.

While (Huang et al. 2017) point out the seven V's of Big Data, we would like to highlight four dimensions characterized by four V's of Big Data, that challenges the traditional solutions of databases, in terms of data acquisition, storage, management and analysis. The first V concerns the Volume of data generated by traditional and new providers. To illustrate the data deluge, let consider three examples of data providers:¹ (i) the massive use of sensors (e.g. 10 Terabyte of data are generated by planes every 30 minutes), (ii) the massive use of social networks (e.g., 340 million tweets per day), (iii) transactions

(Walmart handles more than 1 million customer transactions every hour, which is imported into databases estimated to contain more than 2.5 Peta-bytes of data). The second V is associated to the Variety, where data may come from various data sources, in different formats such as transactions, log data, social network, sensors, etc. from various applications, structured data as database table, semi-structured data such as XML data, unstructured data such as text, images, video streams, audio statement, and more. The third V is about the Velocity, where large amounts of data from transactions with high refresh rate resulting in data streams coming at great speed. The time to act on the basis of these data streams will often be very short. Consequently, the traditional batch processing has to move to real time streaming. The fourth V that got little attention is related to Vocabulary used to describe schemes, data models, semantics, ontologies, taxonomies, and other content- and context-based metadata that describe the data's structure, syntax, content, and provenance.

The traditional database non-functional requirements were mainly focused on (i) improving low-latency query processing to satisfy the needs of end users (e.g., decision makers), (ii) minimizing the maintenance cost of the databases and optimization structures (e.g. indexes and materialized views) and (iii) better usage of the storage cost dedicated to store the optimization structures have been considered. These requirements have been enriched by new ones; mainly motivated by the development of green computing and services delivered by Cloud Computing. This for some years now, the international community regrouping states, governments, associations, and users has been closely involved in climate change by proposing initiatives to limit global warming. The database community has spared no effort propose initiatives in this sense. As a consequence, the reduction of energy has become a new non-functional requirement integrated in the processes of design and the exploitation of database and information systems (Roukh et al. 2017). The Claremont report on database research states the importance of designing power-aware DBMSs that limit energy costs without sacrificing

¹<http://www.economist.com/node/15579717>

✉ Ladjet Bellatreche
bellatreche@ensma.fr
Patrick Valduriez
Patrick.Valduriez@inria.fr
Tadeusz Morzy
tadeusz.morzy@put.poznan.pl

¹ LIAS/ISAE-ENSMA - Poitiers University, Poitiers, France

² INRIA and LIRMM, Montpellier, France

³ Institute of Computing Science, Poznan University of Technology, Poznan, Poland

scalability. This is also echoed in the more recent Beckman report on databases, which considers the energy constrained processing as a challenging issue in Big Data (Abadi et al. 2016). Another non-functional requirement that emerges with the development of Cloud computing is the *pricing* (Pay-as-You-Go) (Toosi et al. 2016). This is because Cloud computing providers offer numerous on-line services based on SLA (Service Level Agreement) between them and their customers.

Considerations above open the door to innovative research directions and challenges for the database research community, yet exploiting two opportunities: (i) computational and storage resource modeling and organization; (ii) Big Data programming models and (iii) processing power. This can be accomplished by means of actual powerful hardware and infrastructures as well as new programming models available at now.

Opportunity 1 The database storage systems are evolving towards decentralized commodity clusters that can scale in terms of capacity, processing power, and network throughput. The efforts that have been deployed to design such systems *share* simultaneously physical resources and data between applications. Cloud computing largely contributed in augmenting sharing capabilities of these systems thanks to their nice characteristics: *elasticity*, *flexibility*, *on-demand storage* and *computing services*.

Opportunity 2 Despite the data deluge and spectacular power of machines, traditional programming languages show their limitations. A new generation of programming models has been proposed, known by *Big Data programming models* (Wu et al. 2017). They represent the style of programming and present the interface paradigm for developers to write big data application programs (Wu et al. 2017). A nice classification of these programming paradigms is given in Wu et al. (2017). The authors distinguished eight classes of programming models: (1) Mapreduce (e.g. Hadoop), (2) functional (e.g. Spark, Flink), (3) SQL-based (e.g. HiveQL, SparkSQL), Actor (e.g., Akka, Storm), (5) statistical and analytical (e.g. R, Mahout), (6) dataflow (e.g. Oozie, Dryad), (7) Bulk Synchronous Parallel (BSP) (e.g. Giraph, Hama) and high-level DSL (e.g. Pig Latin, Linq).

Opportunity 3 Traditionally, storage systems managed databases that were primarily stored on secondary storage and only a small part of the data could fit in main memory. Therefore, disk Input-Output (IO) was the dominating cost factor. Nowadays, it is possible to equip servers with several terabytes of main memory, which allows us to keep databases in main memory to avoid the IO bottleneck (Arnold et al. 2014). Therefore, the performance of

databases became limited by memory access and processing power (Breß et al. 2014). Many heterogeneous devices (e.g. GPU, FPGA, APU) are available and can be used in parallel in order to process database operations, where each processor is optimized for a certain application scenario (Breß et al. 2014, 2016).

The database community, including researchers, industrial and funding organizations has spared no effort to conduct advanced research, by exploiting the achievements and above opportunities, to satisfy the requirements of companies in terms of data management and exploitation.

In the rest of this article, we first discuss some specific research challenges around the databases. We also present, in Section 3, some of the latest development in this research area. We particularly report four research papers addressing different aspects of the database challenges.

2 Databases Research Challenges

The database and information systems technologies raise a good number of research challenges and in this section, we will discuss some of the most important ones, which are related to ADBIS conference, which represents the origin of this special section.

The ADBIS conference is one of concrete examples of these efforts. The ADBIS conference has been widely accepted as a key technology for enterprises and organizations to improve their abilities in data modeling, data management, data exploitation and information systems. The ADBIS conference has attracted the international interest of the research community and is being mentioned in several ranking lists and indexed in several digital libraries (e.g. DBLP, Google Scholar, Microsoft Academic Search).

The first challenge we would like to discuss is the management of evolution of new advanced databases. Methodologies and techniques used for designing advanced databases (data integration systems, data warehouses, data marts, etc.), research developments, and most of the commercially available technologies tacitly assumed that a semantic integration system is static (Bellatreche and Wrembel 2013). In practice, however, this assumption turned out to be false. An advanced database system requires changes among others as the result of: (1) the evolution of data sources, (2) changes of the real world represented in an integration system, (3) the evolution of domain ontologies and knowledge bases (such as DBpedia, FreeBase, Linked Open Data Cloud, Yago, etc.) usually involved in the construction of these databases, (4) new user requirements, and (5) creating simulation scenarios (what-if analysis).

As reported in the literature, structures of data sources change frequently. For example, during the last 4 years, the schema of Wikipedia changed every 9–10 days,

on the average. From our experience, schemas of data sources may change even more frequently. For example, telecommunication data sources changed their schemas every 7–13 days, on the average. Banking data sources are more stable but they changed their schemas every 2–4 weeks, on the average. Changes in the structures of data sources impact all the layers of advanced database systems. Since such changes are frequent, developing solutions and tools for handling them automatically or semi-automatically is of high practical importance.

The *data cleaning*, also called data cleansing or scrubbing represents a serious challenge when integrating data from various sources (Rahm and Do 2000). It aims at detecting and removing errors and *inconsistencies* from data in order to improve the quality of data. Data cleaning is especially required when integrating heterogeneous data sources and should be addressed together with schema-related data transformations. In data warehouses, data cleaning is a major part of the so-called ETL (Extract, Transform, Load) process. The first task of an ETL process is to extract data from multiple data sources, typically into a Data Staging Area (DSA). Once data are available in a DSA, the second phase is to perform data quality checks and transformations in order to make data clean and consistent with the structure of a materialized integration system. Finally, the third phase is to load data into an integration system. The ETL process is error-prone and cause significant downtime; especially, when it deals with a deluge of data issued from various and heterogeneous data sources. The causes of these errors concern mainly availability of data sources, reading, writing, incomplete data, duplicate, system crash or power outage machine hosted ETL, etc. The development of rigorous of testing procedure before the loading task contributes in increasing the quality of data integration systems.

A connected issue to the ETL quality concern its performance. It should be noticed that ETL processes are performed outside DBMS. An ETL process is typically implemented as a work-flow, where various tasks (a.k.a. activities or operations), which process data, are connected by data flows (Ali and Wrembel 2017). Optimizing this work-flow is one of important challenges for ETL research community, especially, in Big Data Era, where various sources are involved in the integration process, with a deluge of data. The recourse to Big Data programming models are highly recommended (Ali and Wrembel 2017).

The fourth challenge is related the vocabulary brought by Big Data. The vocabulary has a great role in designing new applications requiring semantics. Let us consider the example of an emerging topic which is the Cultural Heritage that faces an array of challenges. Scientific researchers, organizations, associations, schools are looking for relevant technologies for accessing, integrating, sharing, annotating,

visualizing, analyzing the mine of cultural collections by considering profiles and preferences of end users. Most cultural information systems today process data based on the syntactic level without leveraging the rich *semantic* structures underlying the content. Moreover, they use multiple thesauri without a formal connection between them. This situation has been identified in the 90's when the need to build a unique interface to access huge collection of data has appeared. During the last decades, *Semantic Web* solutions have been proposed to explicit the semantic of data sources and make their content machine understandable and interoperable. Efforts on integrating the Web Semantic technology in the Cultural Heritage have to be taken, by constructing adequate ontologies and vocabularies (Markhoff et al. 2017).

Finally, another challenge that has to be addressed is the development of cost models evaluating the non-functional requirements. Generally speaking, a cost model (\mathcal{CM}) can be seen as a mathematical function with input parameters and as an output the value of the measured cost in terms of response time, size and/or energy. A \mathcal{CM} has five main roles: (i) it selects the best query plans (Andrès et al. 1995), (ii) it guides algorithms to select optimization structures such as indexes, materialized views, etc. Bellatreche et al. (2000), (iii) it is used to deploy a database in advanced platforms (parallel, Cloud, etc.) Kunjir et al. (2017), (iv) it is used by advisory tools to assist database administrators in various of systems tuning and physical design (Chaudhuri and Narasayya 1998) and (v) self-driving database management systems (Pavlo et al. 2017). A \mathcal{CM} is difficult to develop, since it includes several parameters belonging to databases, platforms, DBMS, queries, devices, etc. Ouared et al. (2016). The evolution of non-functional requirements, databases, processing devises, storage systems, etc. pushes the database community to propose methodologies and guidelines to construct, calibrate and validate \mathcal{CM} s.

3 Papers in this Special Section

We got 14 papers for our special section distributed as follows:

- two papers from the main conference ADBIS (Morzy et al. 2015);
- the best paper of each ADBIS workshop. It should be noticed that MEBIS and GIG Workshops were merged. In total, we got 6 papers from ADBIS workshops;
- six papers from our open call.

Authors of selected papers (from ADBIS main conference and its workshops) were invited to submit an extended version with at least 30% difference in technical content.

These papers were evaluated by at least two reviewers. After a second round of reviews, we finally accepted four papers. Thus, the relative acceptance rate for the papers included in this special section is competitive (28.5%). Needless to say, these four papers represent innovative and high quality research. The topics of these accepted papers are very timely and include: Big Data Applications and Principles, Evolving Business Intelligence Systems, Cultural Heritage Preservation and Enhancement and database evolution management.

We congratulate the authors of these four papers and thank all authors who submitted articles to ADBIS 2014 and our special section. It should be noticed that certain papers used case studies issued from international projects funded either by European Commission or German Research Foundation.

The four selected papers are summarized as follows:

The first paper titled: *Evaluating Queries and Updates on Big XML Documents* by Carlo Sartiani, Nicole Bidoit, Dario Colazzo and Noor Malla presents *Andromeda* – a system able to execute a subcategory of iterative and update XQuery Queries over MapReduce (Sartiani et al. 2018). This subcategory is identified as queries that iterate over forward XPath, and can therefore easily be distributed by partitioning the document according to these paths. The authors described the global architecture, the basic principle and the used algorithms of *Andromeda*. The basic idea of this system is to dynamically and/or statically partition the input data to leverage on the parallelism of a Map/Reduce cluster and to increase the scalability. A great effort on formalization of iterative XQuery queries and updates has to be highlighted. Two partitioning algorithms are given for iterative XQuery queries and updates, respectively. Intensive experiments have been conducted on a multi-tenant cluster composed of a single master machine and 100 slave machines. Two distinct datasets have been used covering iterative and updates queries. The proposal is compared against the existing systems such as BaseX.

The second paper titled: *Dependency Modelling for Inconsistency Management in Digital Preservation - The PERICLES Approach*, by Nikolaos Lagos, Marina Riga, Panagiotis Mitzias, Jean-Yves Vion-Dury, Efstratios Kontopoulos, Simon Waddington, Georgios Meditskos, Pip Laurenson, and Ioannis Kompatsiaris presents an important aspect of the PERICLES project: the Linked Resource Model (LRM). A conceptual model to handle contextual and environmental constraints, focusing on the preservation of cultural heritage is presented (Lagos et al. 2018). In the proposed approach, models provide an abstract representation of essential aspects of the environment. The goal is to model details of dependencies concerning components of the artwork. For instance, a dependency between MS Visio

and JPEG objects. The LRM provides concepts that allow recording when a change is triggered and its impact on other entities. Here the LRM is extended by links to other existing ontologies giving rise to Digital Video Art (DVA), a domain-specific model. Examples of automatic reasoning and handling of inconsistencies are given, based on the use of SPIN. The presented case study concerns the preservation of video art and considers the problem of how technology changes are managed and traced over time. The particularity of this paper is its usage of a domain-independent ontology (LRM), combined with domain-specific ontology, to model changes to video art. It also applied to in detecting inconsistencies. This work presents one of the key outcomes of the PERICLES FP7 project (<http://pericles-project.eu/>).

The third paper entitled *Robust and Simple Database Evolution*, by Kai Herrmann, Hannes Voigt, Jonas Rausch, Andreas Behrend and Wolfgang Lehner presents a domain specific language (DSL), named CoDEL, to support the database evolution process. Apart from the language concepts and syntax, the authors prove that their language is as expressive as the relational algebra. The paper starts from a very important and real supposition that, nowadays, the process of software evolution including updates to meet frequent changes of the user requirements is much better supported than the process of database evolution. The main argument given by the authors of this paper is based on the fact that the database evolution process often presents a major bottleneck in the whole software evolution process (Herrmann et al. 2018). CoDEL is a well-defined relationally complete database evolution language and consequently, it can serve as a reference language for productive implementations of database evolution in DBMSs or as foundation for further support of database evolution. As an example, the authors present semi-automatic variant co-evolution. This work is funded by the German Research Foundation (Deutsche Forschungsgemeinschaft; DFG) within the RoSI research training group (GRK 1907).

The fourth paper titled *ETL Workflow Repairation by Means of Case-Based Reasoning*, by Artur Wojciechowski presents a topic of high interest dealing with the following problem: How to cope with data source evolution in the ETL context? The paper proposes a prototype ETL framework, called E-ETL for repairing ETL workows semi automatically when structural changes in data sources occur (Wojciechowski 2018). The framework is based on a *Case-Based Reasoning* method. It consists of two main algorithms: (1) Case Detection Algorithm and (2) *Best Case Searching Algorithm* for choosing the most appropriate case. A test case scenario is presented for illustrating the approach. An experimental evaluation of the approach with respect to its performance is presented, for 6 different factors influencing the performance.

Acknowledgments We hope readers will find the content of this special section interesting and will inspire them to look further into the challenges that are still ahead before designing and exploiting advanced database and information systems.

The guest editors of this special section wish to express their sincere gratitude to all the authors who submitted their papers to this special section. We are also grateful to the Reviewing Committee for the hard work and the feedback provided to the authors. As guest editors of this special section, we also wish to express our gratitude to the Editors-in-Chief: Professor R. Ramesh and Professor H.R. Rao for the opportunity to edit this special section for our ADBIS conference, their assistance during the special section preparation, and for giving the authors the opportunity to publish their work in Information Systems Frontiers. We would like to mention that it is the first time that ADBIS conference organizes a special section for Information Systems Frontiers Journal, Springer, and this coincides with the first French organization of ADBIS in 2015 in Poitiers City. Last but not least, we wish to thank the Journal's staff for their assistance and suggestions.

References

- Abadi, D., Agrawal, R., Ailamaki, A., Balazinska, M., Bernstein, P.A., Carey, M.J., Widom, J. (2016). The beckman report on database researchs. *Communications of the ACM*, 59(2), 92–99.
- Ali, S.M.F., & Wrembel, R. (2017). From conceptual design to performance optimization of etl workflows: current state of research and open problems. *The VLDB Journal*.
- Andrès, F., Kwakkel, F., Kersten, M.L. (1995). Calibration of a DBMS cost model with the software testpilot. In *CISMOD* (pp. 58–74).
- Arnold, O., Haas, S., Fettweis, G., Schlegel, B., Kissinger, T., Lehner, W. (2014). An application-specific instruction set for accelerating setoriented database primitives. In *Acm sigmod* (pp. 767–778).
- Bellatreche, L., Karlapalem, K., Schneider, M. (2000). On efficient storage space distribution among materialized views and indices in data warehousing environments. In *Acm cism* (pp. 397–404).
- Bellatreche, L., & Wrembel, R. (2013). Special issue on: Evolution and versioning in semantic data integration systems. *Journal of Data Semantics*, 2(2–3), 57–59. Retrieved from <https://doi.org/10.1007/s13740-013-0020-6>.
- Breß, S., Funke, H., Teubner, J. (2016). Robust query processing in coprocessor-accelerated databases. In *Acm sigmod* (pp. 1891–1906).
- Breß, S., Siegmund, N., Heimel, M., Saecker, M., Lauer, T., Bellatreche, L., Saake, G. (2014). Load-aware inter-co-processor parallelism in database query processing. *Data and Knowledge Engineering*, 93, 60–79.
- Chaudhuri, S., & Narasayya, V.R. (1998). Autoadmin what-if index analysis utility. In *Acm sigmod* (pp. 367–378).
- Herrmann, K., Voigt, H., Rausch, J., Behrend, A., Lehner, W. (2018). Robust and simple database evolution. *Information Systems Frontiers*, 20(1). <https://doi.org/10.1007/s10796-016-9730-2>.
- Huang, S.-C., McIntosh, S., Sobolevsky, S., Hung, P.C.K. (2017). Big data analytics and business intelligence in industry. *Information Systems Frontiers*, 19, 1229. <https://doi.org/10.1007/s10796-017-9804-9>.
- Kunjir, M., Fain, B., Munagala, K., Babu, S. (2017). ROBUS: fair cache allocation for data-parallel workloads. In *Acm sigmod* (pp. 219–234).
- Lagos, N., Riga, M., Mitziias, P., Vion-Dury, J.-Y., Kontopoulos, E., Waddington, S., Kompatsiaris, I. (2018). Dependency modelling for inconsistency management in digital preservation - the pericles approach. *Information Systems Frontiers*, 20(1). <https://doi.org/10.1007/s10796-016-9709-z>.
- Markhoff, B., Nguyen, T.B., Niang, C. (2017). When it comes to querying semantic cultural heritage data. In *New trends in databases and information systems: Adbis 2017 short papers and workshops* (pp. 384–394).
- Morzy, T., Valduriez, P., Bellatreche, L. (Eds.) (2015). *Advances in databases and information systems*, Vol. 9282. Berlin: Springer.
- Ouared, A., Ouhammou, Y., Bellatreche, L. (2016). Costdl: A cost models description language for performance metrics in database. In *21st international conference on engineering of complex computer systems, ICECCS* (pp. 187–190).
- Pavlo, A., Angulo, G., Arulraj, J., Lin, H., Lin, J., Ma, L., Zhang, T. (2017). Self-driving database management systems. In *CIDR*.
- Rahm, E., & Do, H.H. (2000). Data cleaning: Problems and current approaches. *IEEE Data Engineering Bulletin*, 23(4), 3–13.
- Roukh, A., Bellatreche, L., Bouarar, S., Boukorca, A. (2017). Eco-physic: Eco-physical design initiative for very large databases. *Information Systems Journal*, 68, 44–63.
- Sartiani, C., Bidoit, N., Colazzo, D., Malla, N. (2018). Evaluating queries and updates on big xml documents. *Information Systems Frontiers*, 20(1). <https://doi.org/10.1007/s10796-017-9744-4>.
- Toosi, A.N., Khodadadi, F., Buyya, R. (2016). SipaaS: Spot instance pricing as a service framework and its implementation in openstack. *Concurrency and Computation: Practice and Experience*, 28(13), 3672–3690.
- Wojciechowski, A. (2018). Etl workflow reparation by means of case-based reasoning. *Information Systems Frontiers*, 20(1). <https://doi.org/10.1007/s10796-016-9732-0>.
- Wu, D., Sakr, S., Zhu, L. (2017). Big data storage and data models. In *Handbook of big data technologies* (pp. 3–29).

Ladjet Bellatreche is a full Professor at National Engineering School for Mechanics and Aeronautics (ISAE-ENSMA), Poitiers, France, where he joined as a faculty member since Sept. 2010. He leads the Data and Model Engineering Team of the Laboratory of Computer Science and Automatic Control for Systems (LIAS). Prior to that, he spent eight years as Assistant and then Associate Professor at Poitiers University, France. He was a Visiting Professor of the Québec en Outaouais, Canada, a Visiting Researcher at the Department of Computer Science, Purdue University, USA and the Department of Computer Science of Hong Kong University of Science and Technology, China (1997-1999). His research interest focuses on Data and Model Management. He has co-authored more than 220 papers and received > 2443 citations (H-index=25). He serves in the Editorial Boards of International Journal of Reasoning-based Intelligent Systems, Inderscience and Scalable Computing and Communications, Journal, Springer. He organized/co-organized numerous international and French Conferences and Workshops (ADBIS, DAWAK, ACM DOLAP, MEDI, IEEE SCC, SOFSEM, WISE, EDA, BDA). He has acted as evaluator for funding agencies in Algeria, France, EU, Czech Republic. He actively contributes in promoting research in Africa and Asia, where he co-supervised several students. He co-found several conferences and workshops (ICT-EurAsia, MEDI, CIIA, OAIS).

Patrick Valduriez is a senior researcher at INRIA, heading the Zenith team in Montpellier. He has also been a professor of Computer Science at University Paris 6 (1999-2002) and a researcher at Microelectronics and Computer Technology Corp. in Austin, Texas (1985-1989). He received his Ph. D. degree and Doctorat d'Etat in CS from University Paris 6 in 1981 and 1985, respectively. His research focuses on big data management, in particular distributed and parallel data management. He has authored and co-authored over 400 technical papers and several textbooks, among which "Principles of Distributed Database Systems". He has been a member of the SIGMOD board, a trustee of the VLDB endowment and an associate editor of several journals, including ACM TODS, the VLDB Journal, Distributed and Parallel Databases, and Internet and Databases. He has served as PC chair of major conferences such as PDIS93, SIGMOD97, VLDB98 (industrial chair), VLDB99 (European chair). He was the general chair of SIGMOD04, EDBT08 and VLDB09. He was the recipient of the 1993 IBM scientific prize in CS in France. He obtained the best paper award at VLDB00. He was awarded the 2014 Inria - Académie des Sciences - Dassault Systems Innovation Prize. He is a Fellow of the ACM.

Tadeusz Morzy is a professor in the Computing Science Department of Poznan University of Technology. He received his M. Sc., Ph. D. and Polish Habilitation from the Technical University of Poznań, Poland. Currently, he is professor of computer science at the Institute of Computing Science of the Technical University of Poznań. He has held visiting positions at the Loyola University, New Orleans in USA, Klagenfurt University in Austria, University La Sapienza in Italy, and Free University Amsterdam. He has authored and coauthored over 100 papers on databases, data mining, and data warehousing. He is co-author of a book on "Concurrency Control in Distributed Database Systems" by North-Holland, and editor and coauthor of "Handbook on Data Management" by Springer. He served as General Chair of the 2nd and 16th ADBIS Conferences (1998, 2012), and has served/serves on numerous program committees of international conferences and workshops. His research interests include data mining, data warehousing, transaction processing in database and data warehouse systems, access methods and query processing for databases, database optimization and performance evaluation.