

Keynote 1

HiFun - A High Level Functional Query Language for Big Data Analytics

by *Nicolas Spyratos*, University of Paris South, France

Abstract

There are many systems developed today for the parallel processing of big datasets. Each of these systems is carefully optimized in accordance with the final application goals and constraints. However, their evolution has resulted in an array of solutions catering to a wide range of diverse application environments. Unfortunately, this has also fragmented the big data solutions that are now adapted to particular types of applications. At the same time, applications have moved towards leveraging multiple paradigms in conjunction, for instance combining real time data and historical data. This has led to a pressing need for solutions that seamlessly and transparently allow practitioners to mix different approaches that can function and provide answers as an all-in-one solution. Based on these observations, the overall objective of our work is to separate clearly the conceptual and the physical level so that one can express analysis tasks as queries at the conceptual level independently of how their evaluation is done at the physical level. To achieve this objective we propose (a) a high level functional query language, called HiFun, in which we can formulate queries and study their properties at the conceptual level and (b) mappings to existing evaluation mechanisms (e.g. Hadoop or SQL) which perform the actual evaluation of queries. In other words, we propose a language which is agnostic of the application environment as well as of the nature, structure and location of data. Based on the abstract definitions, we propose a formal approach to the rewriting of analytic queries and the generation of query execution plans. We demonstrate the practical use of our language by showing how queries in HiFun can be mapped as MapReduce jobs in Hadoop and as group-by queries in SQL. Additionally, we show how our approach can leverage semantics in data in order to improve performance. We emphasize that, although theoretical in nature, our work uses only basic and well known mathematical concepts, namely functions and their basic operations.

Keywords: Big Data; MapReduce; Query Language; Data Analytics

Biography

Nicolas Spyratos (spyratos@lri.fr) received his BEng degree from the National Technical University of Athens, Greece, his M.Sc. degree from the University of Ottawa, Canada, his Ph.D degree from Carleton University, Canada and his "thèse d'état" from the University of Paris South 11, France. He worked as a researcher for Bell-Northern Research in Canada, and for INRIA and the National Research Council (CNRS) in France, prior to joining the University of Paris South as a full professor in 1983, where he was heading the database group from 1985 to 2011. He is currently Professor Emeritus at the University of Paris South, scientific advisor of Japan Science and Technology (JST), and affiliated senior scientist, at the Institute of Computer Science of Crete, in Greece (<http://www.ics.forth.gr/>). His research interests include databases, big data analytics, conceptual modeling and digital libraries. He is the author of over 200 articles in international journals, books and conferences, and has supervised the work of 24 PhD students. He has also served on the program committee of over 100 international conferences, he has participated in over 25 national, European and international research projects and has served as evaluator for the NSF and the European commission.



Keynote 2

Machine Learning and Particle Swarm Optimization Tools for Searching in Big Databases

by Moncef Gabbouj, Tampere University of Technology, Finland

Abstract

The talk deals with a new paradigm for multimedia search based on content. We present an alternative approach to classical search engines for information retrieval, which can be used for Big and generic multimedia repositories. We introduce an incremental evolution scheme within a collective network of (evolutionary) binary classifier (CNBC) framework. The proposed framework addresses the problems of feature/class scalability and achieves high classification and content-based retrieval performances over dynamic image repositories. The secret behind the success of CNBC is a novel design to implement the backbone of CNBC, namely the binary classifier. This is a special neural network, which is optimally designed using the recently developed evolutionary optimization algorithm called multi-dimensional particle swarm optimization. Particle swarm optimization (PSO) is population based stochastic search and optimization process, which was introduced in 1995 by Kennedy and Eberhart. The goal is to converge to the global optimum of some multidimensional fitness function. Two novel techniques, which extend the basic PSO algorithm, are presented. The first algorithm called multi-dimensional PSO (M-D PSO) deals with problems in which the dimension of the solution space is not known a priori. M-D PSO solves such a problem by introducing two interleaved PSO iteration processes, a positional PSO followed by a dimensional PSO in which the dimension of a particle is allowed to vary. In a multidimensional search space where the optimum dimension is unknown, swarm particles can seek both positional and dimensional optima. Most content-based multimedia search engines available today rely heavily on low-level features. However, such features extracted automatically usually lack discrimination power needed for accurate description of the image content and may lead to poor retrieval performance. To address this problem, we propose an evolutionary feature synthesis technique, which seeks for the optimal linear and non-linear operations over optimally selected features so as to synthesize highly discriminative features. The optimality therein is sought through MD-PSO. The synthesized features are applied over only a minority of the original feature vectors and exhibit a major discrimination power between different classes and extensive CBIR experiments show that a significant performance improvement can be achieved.

Biography

Moncef Gabbouj received his BS degree in electrical engineering in 1985 from Oklahoma State University, Stillwater, and his MS and PhD degrees in electrical engineering from Purdue University, West Lafayette, Indiana, in 1986 and 1989, respectively.



Dr. Gabbouj is an Academy Professor with the Academy of Finland since January 2011. He held several visiting professorships at different universities, including The Hong Kong University of Science and Technology, Hong Kong (2012-2013), Purdue University, West Lafayette, Indiana, USA (August-December 2011), the University of Southern California (January-June 2012), and the American University of Sharjah, UAE, (2007-2008). He holds a permanent position of Professor at the Department of Signal Processing, Tampere University of Technology, Tampere, Finland, where he leads the Multimedia Research Group. He was Head of the Department during 2002-2007, and served as Senior Research Fellow of the Academy of Finland in 1997-1998 and 2007-2008. His research interests include multimedia content-based analysis, indexing and retrieval, machine learning, nonlinear signal and image processing and analysis, voice conversion, and video processing and coding.

Dr. Gabbouj is a Fellow of the IEEE, a member of the European Academy and the Finnish Academy of Science and Letters. He is the past Chairman of the DSP Technical Committee of the IEEE Circuits and Systems Society and member of the IEEE Fourier Award for Signal Processing Committee. He was Honorary Guest Professor of Jilin University, China (2005-2010). He served as Distinguished Lecturer for the IEEE Circuits and Systems Society in 2004-2005, and Past-Chairman of the IEEE-EURASIP NSIP (Nonlinear Signal and Image Processing) Board. He was chairman of the Algorithm Group of the EC COST 211quat. He served as associate editor of the *IEEE Transactions on Image Processing*, and was guest editor of *Multimedia Tools and Applications*, and *the European Journal Applied Signal Processing*. He is the past chairman of the IEEE Finland Section, the IEEE Circuits and Systems Society, Technical Committee on Digital Signal Processing, and the IEEE SP/CAS Finland Chapter. He was also (co-)Chairman of BigDataSE 2015, EUVIP 2014, CBMI 2005, WIAMIS 2001 and the TPC Chair of ISCCSP 2012, 2006 and 2004, CBMI 2003, EUSIPCO 2000, NORSIG 1996 and the DSP track co-chair of the 2013, 2012, 2011, and 1996 IEEE ISCAS. He is also member of EURASIP Advisory Board and past member of AdCom. He also served as Publication Chair and Publicity Chair of IEEE ICIP 2005 and IEEE ICASSP 2006, respectively, and the Innovation chair of ICIP 2011. He is a member of IEEE SP and CAS societies. Dr. Gabbouj is member of the Finnish Academy of Science and Letters.

Keynote 3

Digital Innovation for Emerging Economies

By Kris Singh: President & CEO at SRII, Silicon Valley, California, USA (www.thesrii.org)

Abstract

Digital technologies are rapidly transforming both business practices and societies and are integral to the innovation driven economies of the future. While technology is advancing rapidly, organizations and skills advance slowly. Inventing effective organizations and institutions for the digital economy is the grand challenge of our time. Governments will need to facilitate the transition of workers to new types of digital jobs. This talk will discuss key subjects such as: Digital megatrends (Mobile technology, Big data/Business Intelligence, Cloud Computing, Social Media), Building Digital Infrastructure, New Digital Business Model, Building Innovation Ecosystem, Jobs in the Digital Economy, Education & Entrepreneurship, and Industry/Academia/Government partnership

Biography



Kris Singh has been a senior leader in the IT industry for more than 25 years. Kris has worked for major IT companies in the Silicon Valley like IBM Research, AMD, Intel and National Semiconductor. At IBM Almaden Research Kris has been the Strategic Programs Director for IT Services Research and where he created the global IT Service Research and Innovation Institute, SRII (www.thesrii.org). SRII mission is to drive IT Innovation for major sectors of economy and build a better world. SRII members include senior leaders from industry, academia, research, government and key professional organizations at the global level. SRII organizations and initiatives are growing all around the world. Before IBM Research, he has been the Director at AMD CTO Office responsible for Product and Technology roadmap definition working closely with major customers and industry ecosystem partners. Prior to AMD CTO Office, Kris worked for Intel as a Director of the Data Center technology/ product/solution roadmap planning team. At Intel he worked with CIO's around the world to better understand the Data Center "Total Cost of Ownership (TCO)" model and help drive down the cost of IT Services/Solutions. Before Intel, Kris has been with National Semiconductor where he managed various groups like Product/ Technology development, Global Manufacturing & Suppliers Management, and major Customer programs.

While working in the IT industry Kris has also been connected with academia. Kris has been an Industry Fellow in the College of Engineering at UC Berkeley, and an adjunct professor in the Electrical Engineering & Computer Science department at Santa Clara University, CA. He is on the Advisory Board at several major universities helping drive "Innovation in Education system and building Entrepreneurship program". Kris is also on the board/advisory member of several startup companies in Silicon Valley and he is a regular speaker at major forums/conferences around the world.

Keynote 4

Subject-Oriented Business Process Management

by Udo Kannengiesser, Metasonic GmbH Germany

Abstract:

Decentralized processing of information and loose coupling of applications are foundational concepts for web computing and related technologies. On the other hand, centralized control is still widely regarded as the paradigm of choice when modelling and executing business processes on top of these technologies. As a consequence, most process applications are inflexible and difficult to adapt and maintain.

This keynote introduces Subject-oriented Business Process Management (S-BPM), a stakeholder-centric, communication-based approach to process management. In contrast to approaches based on centralized control flow, S-BPM modularizes processes into separate viewpoints, each of which describes the behaviour of a different process participant (or “subject”). A single overall process model emerges from the interconnection of the partial viewpoints via messages. Such a clear separation of concerns, combined with the very simple notation of S-BPM, allows domain experts to readily model their work procedures from their own, “subjective” perspective. Changes in the process can be localized and encapsulated in a single module of the process without necessarily affecting others. This is a major departure from traditional approaches where process models are typically monolithic and therefore difficult to change.

The keynote provides an introduction to the innovative concepts, tools and applications of S-BPM, highlighting its benefits with respect to traditional approaches such as BPMN. In particular, results from an EU FP7 project are shown where S-BPM is used as the backbone of next-generation Internet of Things (IoT) applications in manufacturing. An example application is reported in which wearable human sensors and machine actuators are integrated via S-BPM process models.

Biography:

Udo Kannengiesser is a researcher in business process management with over 15 years of experience in modelling dynamic processes and developing agile process management systems. He has published more than 60 research papers in the fields of business process management, information systems, design science and artificial intelligence. His work on the FBS process framework of design has become a major reference for design researchers worldwide and has been applied in various design disciplines including business process design. He obtained his PhD from the University of Sydney (Australia) and worked as a research scientist at National ICT Australia. He is currently a senior research engineer at Metasonic GmbH (Germany), where he leads research projects related to agile process management.

