

## Report on the 2004 SIGMOD Conference



**Gerhard Weikum**

Max-Planck Institute of Computer Science  
Saarbrücken, Germany  
weikum@mpi-sb.mpg.de



**Patrick Valduriez**

INRIA  
Nantes, France  
Patrick.Valduriez@inria.fr

### Introduction

The 2004 ACM SIGMOD International Conference on Management of Data, held in Paris in the week of June 13-18, was the first SIGMOD ever held outside of America. It had chosen a place that is rich in tradition but also rich in new departures, one of the focal points of the age of enlightenment and the place of the French revolution in 1789. At SIGMOD 2004 no revolution happened; the 10-minute power outage that struck some of the presentations was merely caused by a strike of the electricity utility workers - “unmouvement social” in French. However, the technical program of SIGMOD did indicate that the database research community is well on its way into a broader agenda with new challenges arising from business intelligence and scientific information management as well as network-enabled data-rich applications.

The conference was located at La Maison de la Chimie, which was created in 1934, in the heart of Paris, to foster synergy between science and industry. The conference brought together more than 600 scientists, practitioners, and students from all over the world, with the USA (238), France (108), Germany (51), and Italy (28) having had the

most participants. The program included 22 research sessions with 69 papers, 24 demos, 6 industrial sessions, 4 tutorials, 1 panel, and 2 keynotes, and it was co-located with PODS as usual and with 7 pre- and post-conference workshops. The 69 accepted research papers were selected by the program committee out of 431 submissions, which set a new record for SIGMOD.

### Keynotes

The first keynote, given by Jim Gray from Microsoft Research and entitled “The Revolution in Database Systems Architecture” reviewed and interpreted the achievements of database technology in the last decade and the ever stronger role of database systems in computing and information management in general. Jim emphasized the object-relational “ecosystem” as a wonderful enabler for application-specific extensions to support modern applications such as Web Services for astronomy and other sciences. His very insightful talk also raised a number of challenging research avenues, including probabilistic reasoning about data and query results, stream processing and sensor networks, smart “gizmos” in an ambient intelligence environment, and enhancing database

systems with self-configuring, self-healing, and self-managing capabilities.

The second keynote, given by Ueli Maurer from ETH Zurich and entitled “New Approaches to Database Security”, drew a fascinating picture of what modern cryptography can do and how it could possibly affect database systems and database federations. Ueli, who is one of the leading scientists in the cryptography and security community, first gave a wonderful mini-tutorial on the fundamental notions of unilateral and multi-lateral security and how they apply to databases, database users, and digital objects. After this part which was highly insightful for non-experts in security, Ueli then shifted to technically more complex material and discussed cryptographic protocols for simulating a trusted party in secret sharing and multi-party commitment. This was an eye-opener and gave great stimulation to all database researchers who are working or plan to work on the ever more important issues of data security and privacy.

### **Panel**

SIGMOD 2004 had only one, but very entertaining panel, organized by Mike Franklin and Jennifer Widom. It addressed a crucially important issue that the research community needs to reflect: the exploding numbers of submissions to the field’s leading conferences and the resulting problems in terms of handling the refereeing process. For SIGMOD, the numbers of research paper submissions increased from 240 in the year 2002 and 331 in 2003 to 431 in 2004. The other major database conferences like VLDB and ICDE have similar experiences, with the most recent conferences exceeding 500 submissions. Inevitably this leads to difficulties in managing very large program committees, steering the discussion process among PC members, ensuring the quality of reviews, coping with resubmitted papers that

were rejected elsewhere, and selecting the acceptable submissions in view of a very large batch of “borderline papers”.

The panel took the form of a brainstorming, with very interesting, partly provocative if not outrageous, ideas from the panelists and the audience. A custom-engineered “applause-o-meter” (with a patent held by Mike Franklin) was used to obtain feedback from the audience on the suitability of various suggestions. Among the more serious ideas were author feedback and imposing a subgroup structure within the PC. Among the more adventurous thoughts were incentive mechanisms for submitting fewer papers and working harder on reviews. A very interesting but controversial theme was reviving the role and appreciation of journals.

Regardless of the high entertainment value of the panel, everybody felt that the rapidly increasing submission rates are a serious and very important issue that the major conferences need to address soon. On one hand, it is a very positive indicator that the database research community is in healthy shape and growing. On the other hand, the situation does require rethinking the refereeing process. All major database conferences are working on improving their processes along these lines.

### **Observations on the Technical Program**

Overall the technical program covered an impressive spectrum of topics from evergreens like query processing and optimization to hot topics such as stream mining, sensor networks, or the integration of text and data. Not surprisingly, the most prevalent topics, based on the authors’ specification of areas when submitting their papers, were XML, data mining, query processing, and streams. Interestingly, among these query processing by far had the best ratio of accepted papers to submissions, which may

tell us where our community's specific strengths are. This perception was confirmed by a little analysis of the relative entropy of words in paper titles regarding the SIGMOD conferences of different years. The top words for the 2004 conference were "stream", "XML", and "query". The industrial program emphasized two major issues: Web Services and, again, query processing. Obviously, query processing still holds many unsolved problems and poses new challenges in a rapidly changing world of large-scale federated, heterogeneous, semistructured, and highly dynamic network-centric data.

### **Awards**

As usual, the conference also featured an awards ceremony. The 2004 SIGMOD Innovations Award was given to Ronald Fagin for his pioneering work on relational design theory, the application of logic to database theory, and the retrieval of fuzzy information. Starting with this year, the award is renamed into the SIGMOD Edgar F. Codd Innovation Award to honor the late father of the relational database model. It was a very nice gesture that the award was presented by Sharon Codd, Ted Codd's widow and an accomplished database researcher herself.

The 2004 SIGMOD Contributions Award was given to Surajit Chaudhuri for his great voluntary work on creating and maintaining the Conference Management Tool (CMT) that is being used by all major database conferences for the electronic handling of submissions and refereeing.

The Test of Time Award for the most influential paper from the SIGMOD conference 10 years ago was shared by the authors of two papers: Vassilis Christophides, Serge Abiteboul, Sophie Cluet, and Michel Scholl for their work "From Structured Documents to Novel Query Facilities" on foundations of

semistructured data predating the XML era, and Michael Carey, David DeWitt, Michael Franklin, Nancy Hall, Mark McAuliffe, Jeffrey Naughton, Daniel Schuh, Marvin Solomon, C. K. Tan, Odysseas Tsatalos, Seth White, and Michael Zwilling for their work "Shoring up Persistent Applications" on the innovative SHORE system built at the University of Wisconsin.

Finally, the Best Paper Award for the 2004 SIGMOD Conference was given to Raymond Ng and Yuhua Cai for their paper "Indexing Spatio-Temporal Trajectories with Chebyshev Polynomials".

### **Social event**

Last but not least, the social event took place at Salle Wagram, a listed historical monument, located near the Arc de Triomphe, dating back to the second Empire, and famous for its cultural events and prestigious parties. There, in a café style dinner, the participants could enjoy a large variety of delicate samples of French nouvelle cuisine and fine wines (and excellent Champagne) provided by the caterer Saint Clair. Inspired by the spirit of Salle Wagram (and a jazz band), the most adventurous participants ended up dancing, perpetuating the Parisian tradition of night life.