Research Track Program Chairs' Welcome

The KDD conference has seen remarkable growth since its origins as an IJCAI workshop in Detroit in 1989, evolving into a full-fledged research conference in 1995, underscoring the important role that data mining as a field has played in extracting knowledge and actionable insights from vast troves of data that are being generated in the digital world as well as in the physical world around us. This year, we received a record-high number of 1,036 submissions to the research program, out of which 151 papers were accepted. Overall, the number of submissions has increased 42% over the last year (KDD '13), while the acceptance rate dropped from 17.2% last year to 14.6% this year. This makes KDD 2014 by far the biggest and the most competitive KDD in its 20-year history.

Among academic conferences, the KDD conference typically has more of an emphasis on research motivated by real-world applications. It is important to keep in mind that it is this synergy of research in areas such as algorithms, computational geometry, database, graph theory, machine learning, natural language processing, statistics, visualization, and many others when applied to problems arising in diverse fields, such as the Web, medicine, biology, and marketing, driving our field forward and making it vibrant and fun.

The breadth of topics covered in this year's research program is truly comprehensive and nicely balanced among social and information networks (6.58% of papers), data mining for social good (5.26%), graph mining (5.26%), statistical techniques for big data (5.26%), recommender systems (4.61%), data streams (3.95%), scalable methods (3.95%), Web mining (3.29%), clustering (3.29%), feature selection (3.29%), applications to health care and medicine (3.29%), mining rich data types (2.63%), and many more. In particular, we are happy to observe an increase in the number of papers related to data mining for social good, as well as works on scaling up algorithms to deal with big data. It is extremely encouraging to see that data mining is branching out to many application areas of societal importance: public policy, education, healthcare, medicine, smart cities, the Internet of things, and many others. We hope that this trend of data mining affecting all aspects of our society will definitely continue in the future.

We are also very fortunate to have four world-class keynote speakers this year spanning industry and academia, providing inspirational talks on cutting-edge techniques and issues in big data, automatic knowledge discovery, and decision making in medicine, science, and understanding human behavior.

The process of whittling down the initial 1,036 submissions to the final set of 151 accepted papers required the coordination and time of a large number of willing volunteers. The program committee (PC) consisted of 376 reviewers (PC members) and 46 senior PC members. The selection of the PC members involved quantitative constraints to achieve a well-balanced coverage of expertise matching the expected distribution of the submissions' subject areas, as well as requiring seniority and freshness in the field. In the first review phase, each submitted paper was automatically assigned to three reviewers (after a bidding process). After the first round of reviewing was complete, the papers were assigned to the senior PC members based on their bids. The senior PC members then initiated discussions on many of the papers, e.g., if there was significant divergence in scores among reviewers or if a paper was on the borderline of being accepted. Following the discussion phase, the senior PC members provided a recommendation score and a meta-review for each paper. In the final phase, we (the program chairs) normalized and analyzed all of this information, starting with the obvious accept and reject decisions, and then we gradually focused in more detail on the papers near the borderline, seeking additional reviews and input from the PC and senior PC members where appropriate. The top-rated papers were provided to an awards committee, chaired by Thorsten Joachims and Sunita Sarawagi, who chose the best paper and best student paper award winners.

We conclude with some well-deserved words of thanks to the large supporting cast. We thank all authors for submitting their research to KDD 2014. We are extremely grateful to the PC members and senior PC members for volunteering their time to help in the review process. Diligent peer review is an essential cornerstone for research progress, and we thank you all for the many detailed and conscientious reviews that were generated. Last year's program chairs, Inderjit Dhillon and Yehuda Koren, were generous in providing us with advice and expertise based on lessons learned in 2013. We would also like to convey our sincere thanks to the staff who run the Microsoft CMT system—they were more than willing to answer questions and help us at all times during the submission, review, and decision process. Finally, we would like to thank all of our colleagues on the KDD 2014 organization committee, who were a pleasure to work with and without whom we would not have such a wonderful conference. Let the 20th ACM SIGKDD conference begin—we hope you enjoy it!

Jure Leskovec and Wei Wang
ACM SIGKDD 2014 Program Chairs