

Mining Genetic Interactions in Genome-Wide Association Study

Wei Wang*

*University of North Carolina
weiwang@cs.unc.edu,
<http://www.cs.unc.edu/~weiwang/>

Summary

Advanced biotechnologies have rendered feasible high-throughput data collecting in human and other model organisms. The availability of such data holds promise for dissecting complex biological processes. Making sense of the flood of biological data poses great statistical and computational challenges.

I will discuss the problem of mining gene-gene interactions in high-throughput genetic data. Finding genetic interactions is an important biological problem since many common diseases are caused by joint effects of genes. Previously, it was considered intractable to find genetic interactions in the whole-genome scale due to the enormous search space. The problem was commonly addressed using heuristics which do not guarantee the optimality of the solution. I will show that by utilizing the upper bound of the test statistic and effectively indexing the data, we can dramatically prune the search space and reduce computational burden. Moreover, our algorithms guarantee to find the optimal solution. In addition to handling specific statistical tests, our algorithms can be applied to a wide range of study types by utilizing convexity, a common property of many commonly used statistics.

Bibliography

Wei Wang is a professor in the Department of Computer Science and a member of the Carolina Center for Genomic Sciences at the University of North Carolina at Chapel Hill. She received a MS degree from the State University of New York at Binghamton in 1995 and a PhD degree in Computer Science from the University of California at Los Angeles in 1999. She was a research staff member at the IBM T. J. Watson Research Center between 1999 and 2002. Dr. Wang's research interests include data mining, bioinformatics, and databases. She has filed seven patents, and has published one monograph and more than one hundred research papers in international journals and major peer-reviewed conference proceedings. Dr. Wang received the IBM Invention Achievement Awards in 2000 and 2001. She was the recipient of a UNC Junior Faculty Development Award in 2003 and an NSF Faculty Early Career Development (CAREER) Award in 2005. She was named a Microsoft Research New Faculty Fellow in 2005. She was recently honored with the 2007 Phillip and Ruth Hettleman Prize for Artistic

and Scholarly Achievement at UNC. Dr. Wang is an associate editor of the IEEE Transactions on Knowledge and Data Engineering and ACM Transactions on Knowledge Discovery in Data, and an editorial board member of the International Journal of Data Mining and Bioinformatics. She serves on the program committees of prestigious international conferences such as ACM SIGMOD, ACM SIGKDD, VLDB, ICDE, EDBT, ACM CIKM, IEEE ICDM, and SSDBM.