Preface on the Reprinted Edition

I am very happy to be contacted for the possibility of having my PhD thesis reprinted. Since my recent conference paper "Practical Tera-Scale Walsh-Hadamard Transform" (`http://ieeexplore.ieee.org/document/7821757/`), I've been working on the two projects: 1) Walsh spectrum analysis and cryptographic applications, 2) Distributed RAM computing - design and implementation.

For the first project, I found out that Walsh spectrum analysis, (which I started initially from some interesting FFT *and* WHT experiments on the sampling distributions of E0 and DES in 2014), has close connections and applications with LPN, linear cryptanalysis, calculation of channel capacity in communications, and high-precision numerical computation and optimization.

In particular, if only a very small number of Walsh coefficients occur in the peaks, it is interesting and important to find the magnitudes and their locations. Though LPN can be seen as a better solution to this problem compared with the plain Fast Walsh-Hadamard Transform, it remains an open question to seek a solution that is not based on the birthday problem. As I have recently discussed informally with some of my colleagues and friends, the Johnson-Lindenstrauss Theorem, 1984 (cf. `http://dx.doi.org/10.1002/rsa.10073`) seems to point a promising direction. Here, the main difficulty is that, from the correct paper formula to the computer implementation of algorithms, we seem to need the source generator of Gaussian noises.

In case that a great number of Walsh coefficients occur in the peaks, it might not be a good idea to find the magnitudes and their locations. In cryptography, one actually only wants to judge whether or not the sampling distribution comes from a uniform distribution, given the sample number. Additionally, two other problems are found to be worth studying further, i.e., 1) Fourier spectrum analysis, 2) the time-domain analysis (in order to identify the peak magnitudes and the peak locations without exhaustive queries).

Currently, the problem of Walsh spectrum on sampling distributions is formulated as follows. Assume that the Walsh spectrum of the biased source distribution can be characterized as: 1) except the zero point, only one coefficient is nonzero; 2) except the zero point, more than one coefficient are nonzero and their absolute values are all equal. We want to analyze the Walsh spectrum for these sampling distributions, given the sample number.

For my second project - Distributed RAM computing, the system consisted of multiple compute nodes will try to use their RAMs jointly to accommodate all data loads; the computation tasks are carried out with RAMs which can be seen as connected virtually via the network (e.g., the ethernet network). I'm working on the preliminary report investigating issues on design and implementation of distributed RAM computing. And I consider that the large-capacity high-speed USB3.0 is potentially the best platform for personal computing at this earlier stage.

The new paradigm of so-called distributed RAM computing has rise on the horizon to address the urgent issue of big data explosion. It is known that in-memory computing was popular in the old days, because of the differences between the one-level storage hierarchy programming paradigm and the multi-level storage hierarchy programming paradigm. Still, everyone finds it troublesome to switch from the one-level storage hierarchy programming paradigm to the multi-level storage hierarchy programming paradigm. As a matter of fact, significant efforts have been put forward widely to encourage and emphasize research experiments to be switched out of the in-memory computing. Thus, the external memory computing and/or massively parallel data-intensive computing are possible solutions to the user. Ideally, from the point of view of the user, one still wants to stick with the one-level storage hierarchy programming paradigm for the complicated and time-consuming data processing tasks. And it leaves the system entirely to solve the problem that can *fast, reliably and efficiently* handle the internal movements of the data items.

Though this second project is expected to have many scientific and engineering applications, it is considered that it is better to be commercialized in the near future in order to speed-up its wide adoption. In my preliminary work, I intend to propose practical implementations and extensions to make a nearly-native virtually extended memory system which is transparent to the user. This way, the user can use the traditional one-level storage hierarchy programming paradigm doing the out-of-core computing.

Yi, June 2017