

# 文献综述

## 一、背景介绍

近年来，在工程应用中，求解高阶矩阵的需求日益增长，全矩阵运算脱离了实际的硬件限制，为了满足这一日益增长的需求，同时这些矩阵通常都有着一个特征——非零元远少于零元，稀疏矩阵这门学科便应运而生。在 20 世纪 60 年代研发电子网络的电子工程师们是最早的去利用稀疏性来应用稀疏矩阵进行工程上德计算的。[1] 而在微分方程数值解、线性规划等的有限元分析中，经常出现求解高阶稀疏线性方程组，如利用全矩阵进行存储，则需要  $n^2$  的空间复杂度和  $n^3$  的乘法运算时间复杂度，显然，这种程度的运算量是无法被微型计算机，甚至是工作站所接受的。而利用矩阵的稀疏性，可以有效地减小消耗很多无谓的存储空间以及无谓的计算，在很大的程度上降低了时间和空间复杂度，降低了计算对硬件的需求。

## 二、国内外研究现状

### 1、研究方向及进展

在这些年的发展中，出现了很多的存储方法，比如：对角线存储法、对称矩阵的变带宽存储法、坐标存储法、Elipack-Itpack 存储法、CSR 存储法、Shermans 存储法、超矩阵存储法、动态存储方案等 [2]。

在早期计算机时，串行的算法占到主流，自然，稀疏矩阵的存储方式以及稀疏矩阵的计算都是为了串行算法服务的。但是随着计算机的发展，计算机集群、多核 CPU、GPU 并行等的出现，让并行算法在计算时间上远远地超越了串行算法，为了更好的适应并行计算中的稀疏矩阵计算，出现了许多新的算法以及存储方式。例如对于非结构化的矩阵，基于 CUDA 框架下的 SCOO 形式的 SpMV 算法比利用基于 Cusp 库的 SCOO 具有更高的效率。[y]

而在做稀疏矩阵的计算时，通常都是做一系列的基本运算，如：矩阵转置、矩阵向量乘法、矩阵矩阵乘法、数乘等。为了能得到更好的效率，许多研究者致力于寻找对于这些计算最优的存储结构及计算算法，同时提供了许多类库供科学计算使用，如：Portable, Extensible Toolkit for Scientific Computation (PETSc)、GNU Scientific Library (GSL) 等。

在多年的研究中，各种相关的研究成果页层出不穷，例如，根据 Michele Martone 的研究成果，在对称矩阵乘法，或者转置的矩阵向量乘法中，RSB 格式的迭代算法效率是比较高的 [x]。

而在李佳佳, 张秀霞, 谭光明, 陈明宇的研究中，得到了影响矩阵性能的参数集，可以利用该参数集提取矩阵特征，并输出最优存储格式，供数值解法器和上层应用调用。[a]

为了适用于大数据集计算以及克服现有稀疏矩阵乘法算法低效的问题，郑建华, 朱蓉, 沈玉利提出了一种基于向量线性组合 (VLC) 的矩阵乘法处理模式，同时采用 MapReduce 计算模型实现了基于 VLC 模式的并行矩阵乘法算法。[d]

而在集群计算方面, 负载均衡是不能不提的, 为了更好地发挥集群计算的计算能力, 付朝江博士给出了基于贪婪分配的稀疏矩阵与向量乘的负载均衡的解决方案。[b]

## 2、存在问题

稀疏矩阵的存储没有一种通用的形式, 各种形式都各有利弊 [f], 这也就需要我们对于不同的情况去寻找最合适的存储形式来计算, 虽然李佳佳, 张秀霞, 谭光明, 陈明宇 [a] 的研究得到了一个参数集用以寻找最优存储形式, 但是比较的存储形式比较少, 还有多种形式没有计入考量范围。

## 3、研究展望

为了更好的适应现代的工程计算需求, 计算更大规模的稀疏矩阵, 以及适应新的 CPU 指令集和 GPU 计算框架, 稀疏矩阵的存储形式值得进一步的研究。同时, 伴随着分布式计算的发展, 更为高效的冗余计算机制和任务分配机制也将带给稀疏矩阵计算新的研究方向。

## 三、参考文献

- [1]Yousef Saad.Iterative Methods for Sparse Linear Systems[M].SECOND EDITION.USA:Society for Industrial and Applied Mathematics,2003 年.68.
- [2] 张永杰、孙秦. 稀疏矩阵存储技术 [J]. 长春理工大学学报,2006 年,03 期:38-41.
- [x]Michele Martone.Efficient multithreaded untransposed, transposed or symmetric sparse matrix-vector multiplication with the Recursive Sparse Blocks format[J].Parallel Computing, 2014,40:47-58.
- [y]Hoang-Vu Dang, Bertil Schmidt.CUDA-enabled Sparse Matrix-Vector Multiplication on GPUs using atomic operations[J].Parallel Computing, 2013,Vol.39 (11):737-750.
- [a] 李佳佳, 张秀霞, 谭光明, 陈明宇. 选择稀疏矩阵乘法最优存储格式的研究 [J]. 计算机研究与发展, Journal of Computer Research and Developmen,2014 年,04 期:882-894.
- [b] 付朝江. 基于贪婪分配的稀疏矩阵与向量乘的负载均衡 [J]. 福建工程学院学报,2010 年,01 期:79-82.
- [d] 郑建华, 朱蓉, 沈玉利.Sparse matrix multiplication algorithm based on Map Reduce[J]. 仲恺农业工程学院学报,2013 年,03 期:45-50.
- [f] 张永杰, 孙秦. 稀疏矩阵存储技术 [J]. 长春理工大学学报,2006 年,03 期:38-41.