

专利申报基本信息

专利名称

一种图结构感知的图神经网络硬件加速器

实验室内部编号

ZL-2023-1-00031

管理归口

科技类

实验室排名

1

联合申报证明材料

📄

高并发分布式图计算系统.pdf

2023-01-09 张宇

申请国别

中国(一般)

是否巴黎公约

否

代理机构

北京博思佳知识产权代理有限公司

实验室单独申请

否

主管部门

专利申请费支出单位

之江实验室

维持费支付单位

之江实验室

是否加急

否

加急理由

专利内容简介

本发明提出了一种用于图神经网络推理的新型软硬协同硬件加速器，包含以下内容：一个新型的社区化算法，用于实现高效的运行时并行图重组，并显著提高片上数据的局部性。一个新的硬件加速器架构，它有效地实现了社区化算法，并通过社区化带来的数据局部性，避免了共享邻居之间的冗余聚合。该硬件架构包含三个主要组件：社区探测器，社区任务分配器和社区处理器。其中社区探测器能够在运行时检测出枢纽节点（即具有高度数的节点），找到它们的邻居，然后以这些邻居节点为起点，迭代地探索和确定社区。任务分配器负责将各个社区处理任务分配给合适的处理单元，针对计算量较大的社区任务，它会采用重映射算法将任务拆分，以确保各个处理单元负载均衡。社区处理器通过使用枢纽和社区信息，以细粒度流水线的方式实现了聚合和组合操作。本发明通过硬件加速器，实现在线图重组，从而显著提高数据的局部性，减少图神经网络推理加速过程的冗余计算。

应用前景分析

图卷积神经网络成为近年来的研究热点，很多工作提出了专门针对图卷积神经网络进行优化的硬件加速器。但是现实世界中图具有大尺寸、高稀疏度、数据局部性差，幂律分布等特点，为图卷积网络加速器带来了很多挑战。具体来说图卷积主要分为图聚合、图组合两个步骤，前者是图卷积神经网络的性能瓶颈。图聚合的实现方案有两种：PULL和PUSH。两种方案各有优劣，但是它们都因为邻接矩阵的不规则性，导致在访问特征矩阵或结果矩阵时，数据重用性很差。针对以上挑战，有不少解决方案尝试通过重构邻接矩阵来改善数据局部性。这些方案都假设输入图在处理的过程中是保持不变的，而实际上现实应用中大部分图都是在时刻更新的。少数已有的解决方案，采用在线的方式处理动态图，结果图重构开销大到难以接受，且他们基于软件的图重排算法导致了严重的延迟。

备注说明

专利申报书

📄

专利-一种图结构感知的图神经网络硬件加速器.docx

2023-01-09 张宇

最后操作人

蔡毅

最后操作时间

2023-01-10

专利权人

排名	专利权人
1	之江实验室
2	华中科技大学

专利所属单位

单位名称
智能计算研究院--图计算研究中心

发明人

排名	姓名	性别	类别	职工号	所属单位	邮箱	已发邮件通知	确认状态
1	张宇	男	其它(外部人员)(博士)		华中科技大学	zhyu@hust.edu.cn	是	是
2	卢浩宇	男	其它(外部人员)(硕士研究生在读)		华中科技大学	hylu@hust.edu.cn	是	是
3	赵进	男	内部员工	002409	智能计算研究院--图计算研究中心	zjin@zhejianglab.com	是	是
4	余辉	男	其它(外部人员)(博士研究生毕业)		华中科技大学	huiy@hust.edu.cn	是	是
5	张湛	男	内部员工	002424	智能计算研究院--图计算研究中心	zhanzhang@zhejianglab.com	是	是

专利所属项目

☐ 无依托项目

序号	所属项目	项目类别	项目负责人
1	高并发分布式图计算系统	科研攻关项目（大于100万）	张宇

审核信息

开始

项目负责人

项目经理（成果管

相关部门审核

结束

相关部门审核 | 蔡毅 | 2023-01-10 11:47:08 |

通过

项目经理（成果管理） | 徐扬 | 2023-01-09 22:12:27 |

通过

项目负责人 | 张宇 | 2023-01-09 16:40:51 |

通过

科研人员 | 张宇 | 2023-01-09 16:40:11 |

提交