Graph Analysis Team Introduction
Experience in GraphBLAS and Graph Databases

Semyon Grigorev

Saint Petersburg State University

April 30, 2022

Group Info

- Established in 2012
- Lead: Semyon Grigorev
  - PhD (2016), Associate professor (2016, SPbSU)
  - dblp: https://dblp.org/pid/181/9903.html
  - h-index (scopus): 5
  - s.v.grigoriev@spbu.ru
- PhD students: 2
- Master students: 5
- Bachelor students: 6
- Publications
  - Total: > 30
  - Scopus: 25

Research Area

- High-performance graph analysis
  - GraphBLAS-based algorithms design, implementation and evaluation
  - Portable multi-GPGPU implementation of GraphBALS-like API
  - GraphBLAS API analysis

Research Area

- High-performance graph analysis
  - GraphBLAS-based algorithms design, implementation and evaluation
  - Portable multi-GPGPU implementation of GraphBALS-like API
  - GraphBLAS API analysis
- Formal Language Constrained Path Querying (FLPQ)
  - New algorithms development
  - Complexity analysis
  - New classes of languages investigation
  - High performance algorithms implementation and evaluation

GraphBLAS API

- Graph-matrix duality
- Operations over matrices and vectors
  - Parametrized by semiring-like structures
  - Based on sparse data structures
  - Highly parallel

---

[1] `https://github.com/DrTimothyAldenDavis/GraphBLAS`
[2] `https://github.com/gunrock/graphblast`
[3] `https://gitee.com/CSL-ALP/graphblas`

GraphBLAS API

- Graph-matrix duality
- Operations over matrices and vectors
  - Parametrized by semiring-like structures
  - Based on sparse data structures
  - Highly parallel

- High-performance implementations
  - SuiteSparse:GraphBLAS[1]: pure C
  - GraphBLAST[2]: GPGPU, Cuda C
  - Huawei's GraphBLAS[3]: C++
  - ...

---

[1] https://github.com/DrTimothyAldenDavis/GraphBLAS
[2] https://github.com/gunrock/graphblast
[3] https://gitee.com/CSL-ALP/graphblas

GraphBLAS API

- Graph-matrix duality
- Operations over matrices and vectors
  - Parametrized by semiring-like structures
  - Based on sparse data structures
  - Highly parallel
- More information on GraphBLAS
  - Home page: `https://graphblas.org/`
  - GraphBLAS-related resources: `https://graphblas.org/GraphBLAS-Pointers/`
  - Introduction to GraphBLAS: `http://mit.bme.hu/~szarnyas/grb/graphblas-introduction.pdf`

- High-performance implementations
  - SuiteSparse:GraphBLAS[1]: pure C
  - GraphBLAST[2]: GPGPU, Cuda C
  - Huawei's GraphBLAS[3]: C++
  - ...

---

[1] `https://github.com/DrTimothyAldenDavis/GraphBLAS`
[2] `https://github.com/gunrock/graphblast`
[3] `https://gitee.com/CSL-ALP/graphblas`

GraphBLAS Applications

- BFS-like algorithms
  - BFS: levels, parents, multiple sources
  - SSSP
  - ...
- Graph clustering

- Transitive closure based algorithms
  - APSP
  - ...
- Triangle counting
- ...

GraphBLAS Applications

- BFS-like algorithms
  - ▶ BFS: levels, parents, multiple sources
  - ▶ SSSP
  - ▶ ...
- Graph clustering

- Transitive closure based algorithms
  - ▶ APSP
  - ▶ ...
- Triangle counting
- ...

- LAGraph: collection of GraphBLAS-based algorithms
  - ▶ GitHub: `https://github.com/GraphBLAS/LAGraph`
  - ▶ Latest report: `https://arxiv.org/pdf/2104.01661.pdf`

Formal Language Constrained Path Querying (FLPQ)

- A way to use formal languages to specify constraints on paths
  - Regular path querying (RPQ): regular expressions as constraints
  - Context-free path querying (CFPQ): context-free grammars as constraints

Formal Language Constrained Path Querying (FLPQ)

- A way to use formal languages to specify constraints on paths
  - ▶ Regular path querying (RPQ): regular expressions as constraints
  - ▶ Context-free path querying (CFPQ): context-free grammars as constraints
- Applications
  - ▶ Graph analysis
  - ▶ Interprocedural static code analysis
  - ▶ Graph database querying

Our Results: GraphBLAS Implementation

- Tools
  - SPbLA: library of GPGPU-powered sparse boolean linear algebra operations
  - Spla: sparse linear algebra framework for multi-GPU computations based on OpenCL
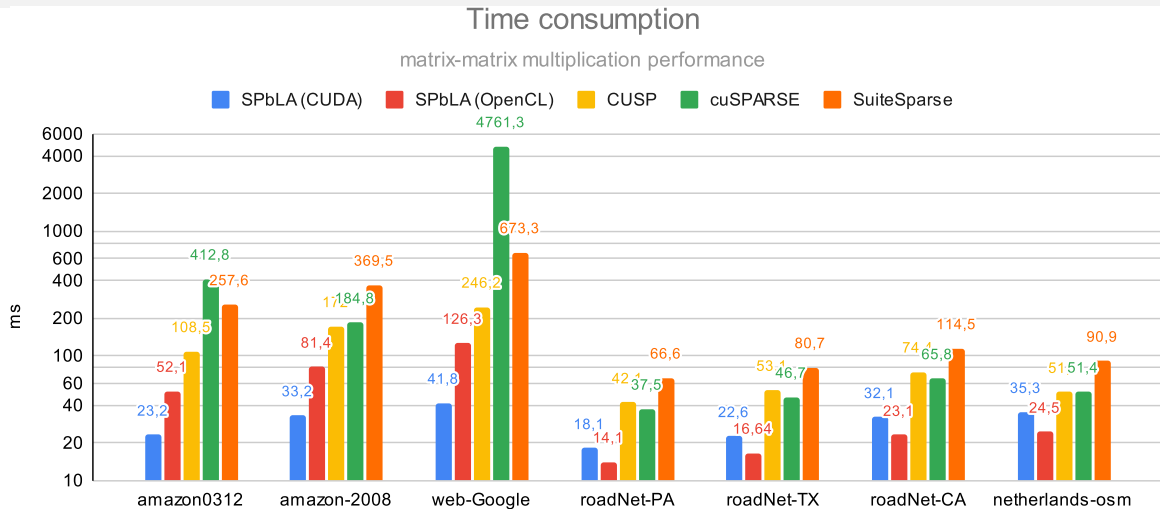
Our Results: GraphBLAS Implementation

- Tools
  - ▶ SPbLA: library of GPGPU-powered sparse boolean linear algebra operations
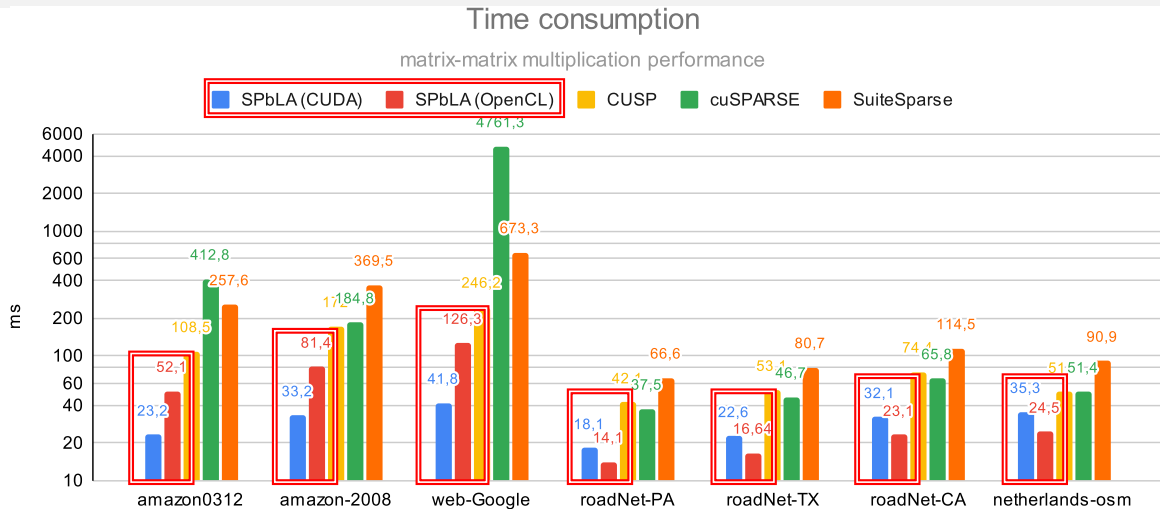 - ▶ Spla: sparse linear algebra framework for multi-GPU computations based on OpenCL
- Papers
  - ▶ SPbLA: The Library of GPGPU-Powered Sparse Boolean Linear Algebra Operations (GrAPL@IPDPS)

Time consumption

matrix-matrix multiplication performance

Legend: ■ SPbLA (CUDA) ■ SPbLA (OpenCL) ■ CUSP ■ cuSPARSE ■ SuiteSparse

[1]More details: https://github.com/JetBrains-Research/spbla#performance

GraphBLAS Implementation Evaluation: SPbLA[1]

Time consumption

matrix-matrix multiplication performance

Legend: SPbLA (CUDA), SPbLA (OpenCL), CUSP, cuSPARSE, SuiteSparse

Categories (x-axis): amazon0312, amazon-2008, web-Google, roadNet-PA, roadNet-TX, roadNet-CA, netherlands-osm

Y-axis: ms

Values (amazon0312): 23,2 / 52,1 / 108,5 / 412,8 / 257,6
Values (amazon-2008): 33,2 / 81,4 / 17_ / 184,8 / 369,5
Values (web-Google): 41,8 / 126,3 / 246,2 / 4761,3 / 673,3
Values (roadNet-PA): 18,1 / 14,1 / 42 / 37,5 / 66,6
Values (roadNet-TX): 22,6 / 16,64 / 53_ / 46,7 / 80,7
Values (roadNet-CA): 32,1 / 23,1 / 74,4 / 65,8 / 114,5
Values (netherlands-osm): 35,3 / 24,5 / 51 / 51,4 / 90,9

GraphBLAS Implementation Evaluation: Spla

Table: Triangles counting algorithms evaluation results. Time in milliseconds (lower is better).

| Dataset | | | Nvidia | | | Intel | |
Name	#V	#E	GR	GB	SP	SS	SP
coAuthorsCiteseer	227.3K	1.6M	2.1	2.0	9.5	17.5	64.9
coPapersDBLP	540.4K	30.4M	5.7	94.4	201.9	543.1	1537.8
roadNet-CA	1.9M	5.5M	34.3	5.8	16.1	47.1	357.6
com-Orkut	3M	234M	218.1	1583.8	2407.4	23731.4	15049.5
cit-Patents	3.7M	16.5M	49.7	52.9	90.6	698.3	684.1
soc-LiveJournal	4.8M	68.9M	69.1	449.6	673.9	4002.6	3823.9

Tools: Gunrock (GR), GraphBLAST (GB), SuiteSparse (SS), Spla (SP)

GraphBLAS Implementation Evaluation: Spla

Table: Triangles counting algorithms evaluation results. Time in milliseconds (lower is better).

| Dataset | | | Nvidia | | | Intel | |
Name	#V	#E	GR	GB	SP	SS	SP
coAuthorsCiteseer	227.3K	1.6M	2.1	2.0	9.5	17.5	64.9
coPapersDBLP	540.4K	30.4M	5.7	94.4	201.9	543.1	1537.8
roadNet-CA	1.9M	5.5M	34.3	5.8	16.1	47.1	357.6
com-Orkut	3M	234M	218.1	1583.8	2407.4	23731.4	15049.0
cit-Patents	3.7M	16.5M	49.7	52.9	90.6	698.3	684.1
soc-LiveJournal	4.8M	68.9M	69.1	449.6	673.9	4002.6	3823.9

Tools: Gunrock (GR), GraphBLAST (GB), SuiteSparse (SS), Spla (SP)

GraphBLAS Implementation Evaluation: Spla

Table: Triangles counting algorithms evaluation results. Time in milliseconds (lower is better).

Dataset			Nvidia			Intel	
Name	#V	#E	GR	GB	SP	SS	SP
coAuthorsCiteseer	227.3K	1.6M	2.1	2.0	9.5	17.5	64.9
coPapersDBLP	540.4K	30.4M	5.7	94.4	201.9	543.1	1537.8
roadNet-CA	1.9M	5.5M	34.3	5.8	16.1	47.1	357.6
com-Orkut	3M	234M	218.1	1583.8	2407.4	23731.4	15049.0
cit-Patents	3.7M	16.5M	49.7	52.9	90.6	698.3	684.1
soc-LiveJournal	4.8M	68.9M	69.1	449.6	673.9	4002.6	3823.9

Tools: Gunrock (GR), GraphBLAST (GB), SuiteSparse (SS), Spla (SP)

GraphBLAS Implementation Evaluation: Spla

Table: Triangles counting algorithms evaluation results. Time in milliseconds (lower is better).

Dataset			Nvidia			Intel	
Name	#V	#E	GR	GB	SP	SS	SP
coAuthorsCiteseer	227.3K	1.6M	2.1	2.0	9.5	17.5	64.9
coPapersDBLP	540.4K	30.4M	5.7	94.4	201.9	543.1	1537.8
roadNet-CA	1.9M	5.5M	34.3	5.8	16.1	47.1	357.6
com-Orkut	3M	234M	218.1	1583.8	2407.4	23731.4	15049.5
cit-Patents	3.7M	16.5M	49.7	52.9	90.6	698.3	684.1
soc-LiveJournal	4.8M	68.9M	69.1	449.6	673.9	4002.6	3823.9

Tools: Gunrock (GR), GraphBLAST (GB), SuiteSparse (SS), Spla (SP)

- Tools
  - ▸ GLL4Graph: CFPQ for Neo4j
  - ▸ CFPQ for RedisGraph

Our Results: Graph Databases Extensions with CFPQ
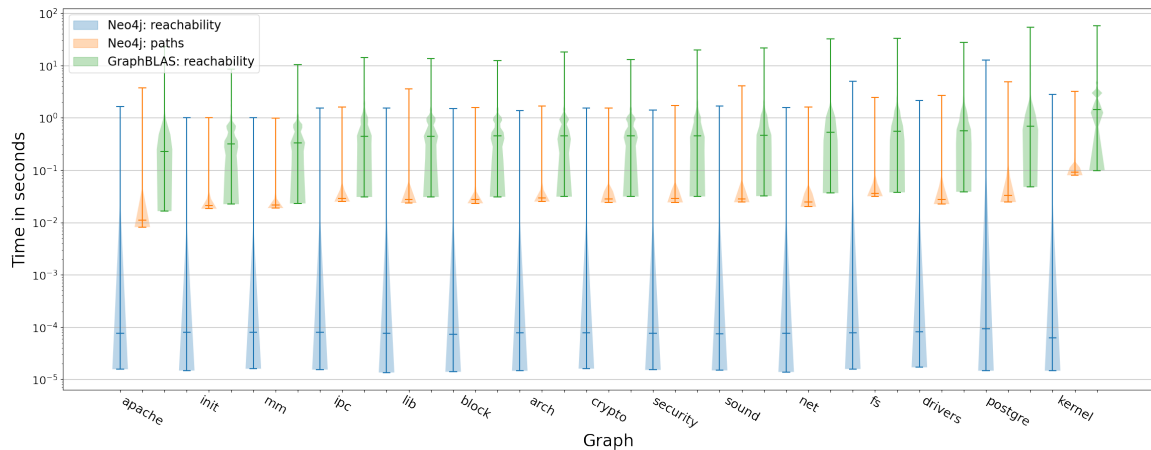
- Tools
  - GLL4Graph: CFPQ for Neo4j
  - CFPQ for RedisGraph
- Papers
  - Multiple-Source Context-Free Path Querying in Terms of Linear Algebra (EDBT, Core A)
  - Context-free path querying by matrix multiplication (GRADES-NDA@SIGMOD)

Graph Databases Extensions with CFPQ: Evaluation

Our Results: FLPQ Algorithms Evaluation

- Tools
  - CFPQ_PyAlgo: set of GraphBLAS-based FLPQ algorithms
    - ✔ GraphBLAS-based CFPQ algorithms
    - ⚙ GraphBLAS-based RPQ algorithms
    - ⚙ GraphBLAS-based MCFG-PQ algorithms

Our Results: FLPQ Algorithms Evaluation

- Tools
  - CFPQ_PyAlgo: set of GraphBLAS-based FLPQ algorithms
    - ✔ GraphBLAS-based CFPQ algorithms
    - ⚙ GraphBLAS-based RPQ algorithms
    - ⚙ GraphBLAS-based MCFG-PQ algorithms
  - CFPQ_Data: dataset for FLPQ algoriths evaluation
    - ✔ Graphs for CFPQ evaluation
    - ⚙ Graphs for RPQ evaluation
    - ⚙ Integration with LDBC Graphalytics

Our Results: FLPQ Algorithms Evaluation

- Tools
  - CFPQ_PyAlgo: set of GraphBLAS-based FLPQ algorithms
    - ✔ GraphBLAS-based CFPQ algorithms
    - ⚙ GraphBLAS-based RPQ algorithms
    - ⚙ GraphBLAS-based MCFG-PQ algorithms
  - CFPQ_Data: dataset for FLPQ algoriths evaluation
    - ✔ Graphs for CFPQ evaluation
    - ⚙ Graphs for RPQ evaluation
    - ⚙ Integration with LDBC Graphalytics
  - ⚙ LDBC Graphalytics extension for evaluation of formal language constrained path querying

Our Results: FLPQ Algorithms Evaluation

- Tools
  - ▶ CFPQ_PyAlgo: set of GraphBLAS-based FLPQ algorithms
    - ✔ GraphBLAS-based CFPQ algorithms
    - ⚙ GraphBLAS-based RPQ algorithms
    - ⚙ GraphBLAS-based MCFG-PQ algorithms
  - ▶ CFPQ_Data: dataset for FLPQ algoriths evaluation
    - ✔ Graphs for CFPQ evaluation
    - ⚙ Graphs for RPQ evaluation
    - ⚙ Integration with LDBC Graphalytics
  - ⚙ LDBC Graphalytics extension for evaluation of formal language constrained path querying
- Papers
  - ▶ Evaluation of the context-free path querying algorithm based on matrix multiplication (GRADES-NDA@SIGMOD)